Developing an Interactive Software Application to Support Young Children's Inference-Making

Kristen L. McMaster, Panayiota Kendeou, Britta Cook Bresina, Susan Slater, Kyle Wagner,

Mary Jane White, Reese Butterfuss, Jasmine Kim, Cristina Umana

Department of Educational Psychology

University of Minnesota

Correspondence concerning this paper should be addressed to Dr. Kristen McMaster,

Department of Educational Psychology, University of Minnesota, 56 East River Road,

Minneapolis, MN 55455; E-mail: mcmas004@umn.edu.

Abstract

We describe the design and development of a technology-based inference-making intervention system that includes a set of interactive learning modules, each of which engages students to (a) view age-appropriate children's videos, (b) learn vocabulary words that are central to main ideas in each video, (c) respond to inferential questions, (d) receive scaffolding and specific feedback for each question, and (e) engage in a set of read-aloud lessons implemented by the classroom teacher and designed to promote transfer of inferencing from non-reading to reading contexts. First, we present the design principles that guided development, drawing on an integrated language comprehension framework. Next, we describe the design process, drawing on a field test of the usability and feasibility of the intervention system. Findings revealed that students and teachers found the system to be usable and helpful for support inference-making, and that it was feasible for classroom use. Then, we provide evidence from a field trial that showed that children who used the intervention system made gains in language comprehension, and that a version with 'offline' questioning (questions asked after viewing videos) was slightly superior to an 'online' version (questions asked during viewing). Finally, we highlight lessons learned that are informing additional development.

Developing an Interactive Software Application to Support Young Children's Inference-Making

Inference making is a skill that is key to children's meaningful engagement in a wide range of literature, as well as to their acquisition of literacy skills needed to succeed in school and later life (Kendeou, McMaster, & Christ, 2016; Snow, 2002). Yet, not all children readily attain this skill, as reflected in national data that indicate, year after year, significant percentages of U.S. students perform below basic proficiency levels in reading (e.g., National Center on Education Statistics, 2017). Whereas many struggling readers' primary difficulties lie in fluent decoding of text, a significant proportion (25%; Helder, Van Leijenhorst, & van den Broek, 2016) can decipher words, but do not successfully comprehend what they read. These comprehension difficulties are often not detected until students reach middle- to late-elementary school grades (Catts, Compton, Tomblin, & Bridges, 2012), when reading achievement gaps become increasingly difficult to close. Thus, prevention and intervention efforts are needed to address comprehension difficulties early on. We propose that inference making should be a significant focus of such efforts.

In this paper, we describe the design and development of a technology-based intervention system to support inference making in young children. The system was designed in the context of multi-tiered systems of support (MTSS), a prevention-oriented approach to assessment and instruction used in many U.S. schools. Within MTSS, children have access to multiple tiers of instructional supports based on need as determined by assessment data. In Tier 1 (also referred to as primary prevention), all students are provided access to high-quality core instruction. In Tier 2 (secondary prevention), those students for whom core instruction is not sufficient receive supplemental intervention designed to target their specific needs. Tier 3 (tertiary prevention) is

reserved for students who require even more intensive, individualized instruction, and might supplement or supplant core instruction, depending on the student's specific needs.

The focus of the current intervention system addresses Tier 1 instruction for 5-year-old (kindergarten) children (Early Language Comprehension Individualized Instruction or ELCII) and Tier 2 intervention for children ages 6 to 7 (Grades 1 and 2) who are at risk for reading failure (Technology-Based Early Language Comprehension Intervention or TeLCI). In the following sections, we first provide a brief overview of the project, which aims to develop an intervention system focusing on children's comprehension of television shows and stories (the larger project addresses fiction and nonfiction; given the focus of this special issue, in this paper we discuss the fiction component). Then, we describe our design principles and process of developing the technology-based intervention system. Last, we describe a transfer component designed to help children transfer inference-making from non-reading to reading contexts.

Project Overview

The overarching criteria guiding development of ELCII and TeLCI were as follows. First, the intervention system must be *theoretically sound*, by adhering to specific instructional design principles, described in detail in the next section. Second, it must be *practically* sound, with evidence of usability and feasibility to be used in classrooms by teachers and students. To be *usable*, students must be able to understand and use the intervention effectively and efficiently; to be *feasible*, teachers must be able to effectively and efficiently implement the intervention in classroom settings. Third, the intervention system must be *empirically sound*, with evidence supporting core components as well as overall promise to improve students' comprehension outcomes. Thus, we examined how to optimize core elements, as well as the promise of the system to improve students' language and reading comprehension outcomes. The design

principles defined how we would know when the intervention was operating as intended, and field testing allowed us to examine usability, feasibility, and promise, providing a feedback loop to revise components that aligned with the design principles.

Design of the Intervention System

In this section, we describe in detail the instructional design principles that guided development of the intervention system, followed by a description of the design process, field-testing, and revision of the system based on student and teacher feedback and empirical data gained along the way. Table 1 provides a list of the design principles, how each was applied to the intervention system, and revisions made based on field-testing.

[Table 1 about here]

Instructional Design Principles

Our instructional design principles draw upon an integrated framework, the Inferential Language Comprehension (*iLC*) Framework (recently articulated in Kendeou, McMaster, et al., in press), which builds on major findings of cognitive, developmental, and language research. The *iLC* framework suggests that inference skills should develop as follows.

Inference skills should transfer across media. There is strong evidence in the literature that inference making is a general skill not specific to reading (Kendeou et al., 2008; Kendeou, van den Broek, White, & Lynch, 2009; Lorch & Sanchez, 1997; Lynch & van den Broek, 2007; Lynch et al., 2008; van den Broek et al., 1996, 2013) and that it can transfer across different media. Indeed, understanding texts presented using different media (visually, aurally, written) involves many of the same cognitive processes (Cohn, 2018; Gernsbacher, 1990; Kim, 2016; Kintsch, 1998; Magliano, Loschky, Clinton, & Larson, 2013). Further, the same structural factors (e.g., number of causal connections, explicitness of goals, event boundaries) predict what

individuals remember from visual, aural, or written narratives (e.g., Lorch & Sanchez, 1997; Lynch et al., 2008;; van den Broek, Helder, & Van Leijenhorst, 2013). In addition, several other processes shared across media can facilitate transfer, such as semantic processing and event segmentation (Cohn & Bender, 2017; Magliano et al., 2013; Magliano, Larson, Higgs, & Loschky, 2016; Zacks, Speer, & Reynolds, 2009).

Conceptualizing inferencing as a 'general skill' that can be developed and transferred across media allows for the use of different media to train inferences in young students independent of decoding skills. Most important, because spontaneous transfer in educational settings is rather difficult (Barnett & Ceci, 2002), it is important to strategically facilitate transfer by training inference skills in different media (for example, initially in visual and subsequently in aural contexts). Thus, our first key design principle is that *inference skills can be trained in one medium (i.e., watching children's videos) and should transfer to another medium (i.e., listening to text read aloud).*

Making inferences requires background knowledge. Researchers have demonstrated that background knowledge fosters and enhances comprehension (e.g., McNamara, Kintsch, Songer, & Kintsch, 1996; Recht & Leslie, 1988), and is critical for inference making (e.g., Cain, Oakhill, Barnes, & Bryant, 2001; McNamara & Kintsch, 1996). The essential role of background knowledge in inference making can be illustrated by examining the two-stage process one undergoes when generating an inference. First, when one encounters information (e.g., while watching a video or reading a text), the current information being presented activates information previously acquired from earlier in the video, text, or from background knowledge. Second, the current information becomes integrated with the prior information, resulting in an inference (Oakhill, 1984). A lack of relevant background knowledge needed to make an inference makes it

difficult to bridge gaps in cohesion, resulting in decreased comprehension (McNamara & Kintsch, 1996). Thus, our second key design principle is that ***inference training should include activation of background knowledge needed for comprehension.*** In our intervention, we support background knowledge by pre-teaching key vocabulary words and concepts needed to comprehend the children's videos.

Questioning should improve inference-making. Numerous examples of question-based interventions that focus specifically on inference generation during reading suggest that questioning can improve inference making (see Elleman, 2017; McMaster & Espin, 2017 for reviews). We theorize that questions prompt inferences because they: (a) cue activation of relevant information, and (b) facilitate integration of that information needed to answer the question. Thus, our third design principle is that ***questioning should facilitate activation and integration of information needed to make inferences*** that support comprehension.

The timing of questioning should be optimized. Whereas questioning interventions generally have been shown to promote inference-making, the theoretical and empirical literature offers limited guidance as to *when* questioning should occur; specifically, whether questions should be asked *online* (i.e., *during* watching of videos) or *offline* (i.e., *after* watching the videos) to support children's overall inference making. Online questions may improve comprehension because they prompt learners to make inferences at the exact points when the inference is needed for comprehension (McMaster & Espin, 2017; Authors, 2012); however, interrupting learners may be detrimental, especially for struggling comprehenders (Goldman, 2004). 'Offline' questions do not impose any interruption in comprehension, but they require that learners have constructed a coherent mental representation during encoding (i.e., as they process incoming information) which they can draw from to answer questions. Thus, the fourth design principle is

that *timing of questioning should be optimized*; this optimal timing needs to be established empirically given limited guidance from the literature.

   Scaffolding and feedback should support inference making. Questioning is often combined with scaffolding, such as prompts (Golke, Dörfler, & Artelt, 2015) and corrective feedback (Kluger & DeNisi, 1996; Sadler, 1989), such as explanations of both correct and incorrect responses (Hattie & Timperley, 2007; Kulik & Kulik, 1988). Questioning with scaffolding and feedback are core ingredients that have potential to support inference making and its acquisition, especially if the scaffolding and feedback are delivered in immediate proximity (before or after) to a student's response to a question (Butler & Roediger, 2007; Authors, 2014; 2015). Thus, the fifth instructional design principle is that *scaffolding and feedback should support inference making* in young students.

   Inference training should provide multiple opportunities to respond. Acquiring and learning new skills typically involves some form of repetition (Cooper & Pantle, 1967), such that learners having multiple opportunities to respond and practice target skills (Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006). Thus, the sixth design principle is that *students should have multiple opportunities to respond* to questioning.

   Inference training should be carefully sequenced. As mentioned earlier, drawing an inference is conceptualized as a two-stage process that involves the activation and integration of information (Oakhill, 1984). Activation and integration are asynchronous, parallel processes with the onset of activation preceding the onset of integration. Asking questions can prompt inferences because they cue activation of relevant information, and facilitate integration of that information to answer the question. Figure 1 illustrates this process: in the context of different media (e.g., visual, aural, or written), information needed for comprehension is activated by

inferential questioning, and further integrated through scaffolding and feedback. If questioning successfully activates information needed to draw an inference and that information gets integrated, then drawing the inference is successful and is validated by explanatory feedback. If drawing an inference is not successful, either because of failures to activate or integrate the needed information, then explanatory feedback is provided along with scaffolding designed to facilitate activation and integration of the needed information.

Note that Figure 1, Panel B shows the final sequence that we adopted for the intervention. Initially, our thinking was that scaffolding should precede questioning (Figure 1, Panel A), such that the sequence was scaffolding—question—feedback. We initially thought this approach would ensure that students would have all information needed to draw an inference activated before the question was asked. In our description of the design process (below), we reveal why we came to the conclusion that the sequence in Panel B is superior to that in Panel A.

[Figure 1 about here]

Design Process

With the above principles as a guide, we set out to develop an interactive software application that trains inference making in young students by (a) engaging students in learning to *make inferences in non-reading contexts* (i.e., while watching children's television shows via videos), (b) include *multiple opportunities to draw inferences* needed to form a coherent mental representation of each video, (c) incorporate *inferential questioning* that prompts inference generation at key points in the videos, (d) provide *immediate scaffolding and feedback* tailored to students' answer selections, and (e) provide *explicit teaching to support transfer* of inference-making skills in children's videos to text read aloud using authentic children's literature.

Drawing on the project criteria and design principles described above, we conceptualized the intervention system as a software application featuring a set of interactive learning modules, each of which engages students to: (a) view age-appropriate children's videos (Principle 1), (b) learn key *vocabulary* words and concepts to provide background knowledge central to the main ideas in each video (Principle 2), (c) respond interactively with the application to *inferential questions* during (online) or after (offline) viewing the video (Principles 3 and 4), (d) receive *scaffolding* and specific *feedback* for each inferential question (Principle 5) with multiple opportunities to respond (Principle 6), (e) during each carefully-sequenced instructional module (Principle 7). Moreover, students would receive *explicit transfer instruction* from the classroom teacher in a read-aloud context (back to Principle 1); we return to this point in a later section.

We followed a six-step iterative process across three years to develop and refine this system. In this section, we describe this process, which was first implemented to develop and test TeLCI as a supplemental Tier 2 intervention for Grades 1 and 2, and then was further refined to develop and test ELCII as a whole-class Tier 1 instructional program for kindergarten. Because the process was largely the same for both TeLCI and ELCII, we focus primarily on TeLCI to provide illustrations of our work. Later, we discuss how we applied lessons learned from TeLCI development to the development of ELCII.

Step 1: Initial module development. In Year 1, we developed the initial TeLCI concept, designed a pedagogical agent (interactive technology-based tutor) to facilitate students' learning, and developed videos with vocabulary, scaffolding, questions, and feedback.

Initial concept. Development of the initial TeLCI concept began in consultation with literacy experts, teachers, and parents. In small focus groups, we presented our idea in very general terms, and asked for reactions both to the overall idea as well as to how specific features

might look and function. For example, educators and parents helped us brainstorm about what form the pedagogical agent would take, identify options for age- and culturally-appropriate children's television shows that would be unfamiliar to students (to support inference-making of novel content), and consider how the system would fit with early elementary curriculum and instructional programming. Teachers emphasized the importance of aligning the content of the application with core literacy standards, and provided guidance regarding ways they envisioned such an application could be integrated into instructional routines, including ideal duration (15 to 20 min) and frequency (2 to 3 times per week).

     Pedagogical agent. To determine the salient characteristics of our TeLCI pedagogical agent, we reviewed the literature to glean characteristics of effective pedagogical agents. We based our workflow and decision-making process on the Pedagogical Agents--Levels of Design (PALD) model (Heidig & Clarebout 2011), which guided our decision-making from *global* levels (e.g., would TeLCI be human or animal?) to *detailed* levels (e.g., what age and gender would TeLCI be; what would he or she wear?). We decided on a non-human (Choi & Clark 2006), peer mentor (Domagk 2010), who could act as a friendly and knowledgeable tutor. We settled on these broad characteristics, and then turned to more specifics.

     We worked with an artist to create the TeLCI pedagogical agent, with a balance between aesthetics and our literature review findings. Eventually, TeLCI emerged as a gender-neutral, purple alien who seemed to be about the same age as our target audience, and who was smart, helpful, and nonjudgmental. When we settled on TeLCI's overall look, the artist designed several different TeLCI positions (e.g., pointing in the direction of vocabulary words or answer selections) and emotions (e.g., excited when a student got an answer correct; puzzled when a

student selected an incorrect answer). TeLCI's voice was provided by a recording artist, who personified TeLCI with a friendly, sing-songy voice.

Videos with questioning, scaffolding, and feedback. Videos (based on television shows for children) for TeLCI modules were selected based on the following criteria: (a) age- and cultural-appropriateness (i.e., would appeal to children from a wide variety of cultural backgrounds) of narratives, (b) level of familiarity students would likely have with the video and its characters (with no familiarity being ideal, to avoid variation in participants' exposure that might confound with their recall and understanding of the videos), (c) continuity in the video passages that were available (with a series of related videos being preferred to unrelated videos), (d) availability of the video passages, (e) potential of the video passages to be edited, and (f) appropriateness of the video passage goal structure--in other words, there had to be a clear goal-action-outcome in the story that was portrayed to be consistent with a focus on narrative structure. An Australian animated television series called *The Adventures of Blinky Bill*, based on a children's book series by Dorothy Wall, was selected for use in TeLCI fiction modules.

To prepare the video passages for use in TeLCI, we modified them to be about 5 min in duration while maintaining a coherent story line with the potential to ask several inferential questions. To determine which content to keep and what to cut in each video, we identified the most important goal of the story and the most salient actions and outcomes related to that goal, allowing us to remove anything unrelated to the goal-action-outcome sequence. The video needed to maintain a clear beginning, middle, and end and make logical sense after editing. During this process, we noted--and sometimes created--gaps in coherence, to create opportunities for students to generate inferences. After editing a video into a 5-min episode, research team-

members not involved in the editing re-watched the video to ensure that a cohesive storyline remained that allowed for inferential questions to be asked during the module.

After a video was successfully modified, we developed five inferential questions per video (in multiple choice format), along with accompanying scaffolding and feedback (see Figure 2 for an example). Five questions seemed to be the maximum number we could ask within a 5-min video that would draw students' attention to the most important content. First, we drafted inference questions and correct answers. These questions were spaced throughout the video, were not repetitive in meaning, and most important, required students to generate inferences with information presented within the video (as opposed to requiring background knowledge) as much as possible. This information was typically related to the goal structure. Second, two scaffolds were identified for each inference question. Scaffolds were intended to *activate* necessary information from the video that students would need to *integrate* in order to make an inference and correctly answer the question. These scaffolds were designed to remind the student of important parts of the video by replaying the corresponding, brief clip from the video. Third, we wrote distractor items for each question. Distractors had to be plausible, of approximately equal syllable length to the correct answer, clearly written, and presented in alphabetical order. Last, we developed feedback for each potential answer selection (correct answer and distractors) for the multiple choice options. The components of the feedback were based on previous intervention work (McMaster et al., 2015). Feedback involved TeLCI kindly telling the student whether or not their answer was correct and why. For example, if the student selected an incorrect option, TeLCI explained why the selected answer was incorrect and which answer was correct, with a brief explanation for the correct answer.

[Figure 2 about here]

Instructional sequence. In addition to developing the videos and questions, we outlined a structure and sequence that each module would follow, with a script for TeLCI to deliver each component. The sequence always began with a greeting from TeLCI, welcoming the student and providing a reminder that the goal was to watch a video and make inferences. In each module, TeLCI reminded the student that an inference is "when we connect something we see, hear, or read to something we already know to make a brand new idea."

Following this brief greeting and reminder, TeLCI introduced the student to three vocabulary words that would be important to understanding the video. Each word was presented by showing a picture that clearly depicted the word (using screenshots from the actual video when possible), along with the word printed clearly below the picture (see Figure 3). TeLCI introduced the word by saying, "The first word is *skipper*. Whisper *skipper* with me: *Skipper*." Then, TeLCI provided a child-friendly definition of the word (e.g., "A *skipper* is the one in charge of steering the ship"), and explained its relevance to the video (e.g., "In the show we're about to watch, the *skipper* is driving the boat that the Circus Brothers are taking to try to meet up with Blinky Bill and his friends"). After introducing all words, TeLCI reminded the student to listen for those words and to be ready to make inferences, and then said, "Let's watch the show!"

[Figure 3 about here]

In the *online* version of TeLCI, the video played up to the point where the first inference was needed; then, TeLCI stopped the video and indicated it was time to make an inference. TeLCI provided scaffolding (illustrated in Figure 2) to prime the student to answer the question, and then asked the question. All together, the questioning sequence went as follows: TeLCI (a) replayed the scaffolding videos, (b) asked the inferential question and provided the answer options, and (c) after the student answered the question, gave targeted feedback that was matched

to the student's specific answer. The *offline* version followed the same sequence, except that the entire video played to the end, and then TeLCI presented each question, one after the other.

We also designed an introductory module in which TeLCI provided a purpose for the modules (to make inferences to help students understand the shows that they were about to watch) and taught what it means to make an inference. To teach about making inferences, TeLCI provided the definition ("when we connect something we see, hear, or read to something we already know to make a brand new idea"). Then, TeLCI gave an example: "If your friend told you that she was going to her cousin's house where there would be friends, cake, and presents, you could *infer* that your friend was going to a birthday party." This example was accompanied by a picture of a birthday party that appeared after TeLCI described the scenario.

Step 2: Initial software application development. After outlining the overall TeLCI concept, identifying video examples, and creating the initial instructional sequence for the modules, we worked with a technology team to create a specification document detailing the functionality of the application, including screenshots, flow charts, and descriptive text. Once we agreed on the overall flow and functionality of the application, the technology team guided us in creating templates for us to insert content that would be used by programmers to create the application. When the initial prototype was completed, the research team reviewed it and made additional adjustments in direct collaboration with the software engineer.

Step 3: Merging learning modules and assessments with the application. We then developed four learning modules (two fiction and two nonfiction). Two versions of each module were created: one with online and one with offline questioning. Both versions were submitted to quality control testing to ensure they were operating as intended, and that the only difference between versions was the timing of questioning.

We also developed two fiction assessments that were almost identical in content and form to the modules. The assessments included TeLCI as the pedagogical agent who asked inferential questions, but *excluded* any inference or vocabulary instruction, scaffolding, or feedback. The purpose of these assessments was to serve as a proximal measure of TeLCI's efficacy in later studies. The assessment content was also entered into the templates that the software team used to merge into the application. As with the learning modules, we created both online and offline versions of the assessments.

Step 4: Year 1 field test. When the first four modules and two fiction assessments were completed and merged with the software application, we conducted a small field test to examine the *usability* and *feasibility* of the TeLCI software application under the two questioning conditions (online and offline). This field test was used to inform further refinement and additional development of TeLCI modules.

Participants. Participants were from two school districts in a large Midwestern city. Eight teachers who provided reading intervention to children in Grades 1-2 volunteered in Fall of Year 1 to participate in the Field Test. In total, 91 students with parental consent participated (grade 1 $n = 71$, grade 2 $n = 20$; see Table 2 for demographic data).

[Table 2 about here]

Measures. We developed several measures of TeLCI's usability and feasibility, including student observations, student and teacher questionnaires, and a teacher focus group protocol.

Two observation forms were developed to observe students completing the TeLCI modules: a group observation form and an individual observation form. The group form included observations pertinent to incidents affecting the whole group of students (e.g., internet problems, delays), while the individual form gathered anecdotal information regarding specific children

(e.g., quality of product, level of interest). On the group form, the observer recorded any

questions or comments that students made about the modules while they were working (e.g., how

to navigate the module, the meaning of words, and so on). Students' apparent engagement and

frustration, as assessed by the observer, was recorded using a 5-point Likert scale (1=low, 5=

high). Technical issues and any additional notes or observations were also recorded. The

individual form was used to focus on individual students to gather more in-depth information

about the TeLCI experience. The observer documented environmental factors (time of day, start

and end times of module, classroom activity, and any interruptions), functionality and usability

of TeLCI (e.g., technical glitches), and the student's interaction and engagement with TeLCI.

After finishing the last TeLCI module, students completed a questionnaire (presented

orally by a researcher) about what they learned from TeLCI and how helpful they thought TeLCI

was in supporting their learning. The questionnaire included questions about the extent to which

students found various features of TeLCI to be in helping them learn, using a three-point rating

scale (1 = very helpful, 2 = kind of helpful, 3 = not very helpful), with follow-up questions for

students to elaborate on their responses. In addition, the questionnaire included open-ended

questions asking what students thought TeLCI's purpose was and what an inference is.

Teachers ($n = 5$) were asked to complete one entire TeLCI module. Then, they completed

a 25-item questionnaire about the academic, cultural, and age-appropriateness of the intervention

(e.g., the appropriateness of materials for a range of learners, what types of students they felt

would benefit most from the intervention, appropriateness of each intervention component), the

usability and feasibility of the intervention (e.g., what the ideal amount of time for each TeLCI

intervention module would be, how the teachers envisioned using the intervention in their

instruction, what data the teachers would like shared with them regarding student performance during the intervention), and what additional features the teachers would like to see.

In focus groups, participating teachers were asked to respond to questions about the TeLCI module, including what features they liked the most and the least, what (if anything) they would add or change about TeLCI, what they thought about the TeLCI character and emotions, and what amount of supervision would be required on their part to use TeLCI in their classroom.

Data collection procedures. In March through May of Year 1, student participants completed two TeLCI fiction assessments (one online, one offline; counterbalanced) and two TeLCI modules (counterbalanced such that each student completed an online fiction, an offline fiction, an online nonfiction, and an offline nonfiction module). Students worked in small groups of 3 to 4 students outside of the classroom, and were supervised by one or two graduate research assistants. Each student completed assessments and modules using a Chromebook tablet and headphones. During this time, the research assistants observed the students and completed the group observation form. When possible, each research assistant randomly selected individual students and completed the individual observation form.

Following student data collection, teachers participated in one of two focus group sessions held at their respective school buildings. The focus group sessions were led by the two principal investigators (co-PIs). Prior to the focus group discussion, each teacher was provided with a Chromebook and asked to complete an entire TeLCI fiction module, followed by completing the teacher questionnaire via Qualtrics. When all teachers had finished this process, the co-PIs led a debriefing discussion (using the focus group questions described above) about the teachers' experience with TeLCI (based on their own experience in using the application, as well as their observations of their students' experiences during the field test).

   Data analysis. Quantitative data from teacher and student questionnaires were aggregated and summarized. Qualitative data from all data sources were coded using a content analysis approach (Krippendorff, 2003) to determine patterns and themes across data sources. We used deductive analysis, starting with categories on the questionnaires, observations, and focus group questions, but also sought to discover emergent categories from the data. The final set of converging feedback from the various sources was synthesized into a checklist used to revise/refine the intervention.

   Field test results. Table 3 summarizes the categories derived from observation notes and their definitions, along with specific notes about what was observed; these notes were later used to inform revisions of TeLCI (described in the next section). Table 4 summarizes students' questionnaire ratings of the extent to which TeLCI features helped them learn while using the software. Below we summarize additional student and teacher feedback.

                                 [Tables 3 and 4 about here]

   Student feedback. Overall, the majority of students reported that they found TeLCI to be *"very helpful"* or *"kind of helpful."* In their comments, students indicated that they liked learning new words at the beginning of each module. The students had mixed feelings about TeLCI stopping to re-watch parts of the video and make inferences. Some of them said they did not like when the video stopped and wanted to keep "reading." Others liked it because they could recall the story better. Regarding the scaffolding, about half of the respondents liked it because it "was good" and helped them remember parts they might have missed. The other half did not like it because it interrupted the video or they did not think they needed help. The offline scaffolding and questioning had slightly better reception, because it did not stop the video. One student commented that it "save[d] all the learning for the end."

Most of the students liked that TeLCI provided feedback. They agreed that telling

them the right and wrong answers was helpful because it helped their comprehension (though

a few of the students thought TeLCI was "rude" and "mean" when correcting their answers).

Regarding TeLCI giving instructions on how to rewind to re-watch important parts and use

the buttons, most students found it useful, although a few said they did not need the

instructions because they already knew what to do.

More than 60% of student respondents reported that they did *not* know or remember

the purpose of using TeLCI. Some said TeLCI's goal was to help you learn new things, while

others mentioned watching stories about Blinky Bill. Only a few students stated that the goal

was to learn how to make inferences. When asked to define "inference," almost 50% said they

forgot what it meant. For some, inferencing involved putting two things (e.g., ideas or

questions) together to make another "big" thing, or to learn "new words and stuff." One

student explained that an inference is "when you are taking two things to think about and you

can make [them] into one big thing you are thinking about." A few students were confused by

the example provided in the introduction--they thought that an inference was a birthday party!

Teacher feedback. After completing an entire TeLCI fiction module, all teachers

deemed the material to be generally appropriate for their students. They liked the vocabulary

presentation, and TeLCI's features, such as different stances to show emotions. They

reiterated (from initial input prior to development) that the ideal length and frequency of the

modules would be 15-20 min (as opposed to the 25-30 min that the initial modules took), 2 to

3 times per week**.** Regarding the application of TeLCI within their instruction, most teachers

considered that the best way to use the software would be during independent work time.

Several features of the modules were considered either *extremely* or *somewhat*

appropriate: the vocabulary lesson, the scaffolding that allowed the students to re-watch

specific segments, the questions for helping the students understand the show, TeLCI's

feedback for the correct responses, and TeLCI's feedback for incorrect responses. All teachers

were interested in being able to access data from students' responses to TeLCI questions,

including the total correct responses and whether scaffolding was needed, as well as the

breakdown of responses by questions and the type of errors made.

Teachers expressed different opinions regarding who would benefit from TeLCI.

Some mentioned students with high comprehension skills and background knowledge. Others

mentioned students who lose interest in reading due to their lack of decoding skills; students

who struggle with low comprehension skills, or language skills, or both; or all students except

English learners who were newcomers (those new to the country within the past few months).

Several teachers also mentioned that, when a student selects an incorrect answer on the first

try, the student should have a second chance to respond.

Step 5: Incorporate field test findings into further development. Findings from the

field test helped us identify several key problems related to our criteria to be *theoretically sound*

(based on instructional design principles) and *practically sound* (based on evidence of

feasibility/usability) and solutions for revision, as follows:

Students did not seem to learn or recall the meaning of 'inference.' This issue relates to

Instructional Design Principle 1, that inference-making is a general skill that can be trained in

different media. We aimed to help students understand what an inference is and that they can

make inferences in a variety of contexts. Our solution was to improve initial instruction about

inferencing and then reinforce its meaning and importance throughout the modules. First, we

made the birthday party example more explicit by clarifying that it was just one example of an

inference, and used three thought bubbles to show how TeLCI was thinking about each part of the story (a picture of friends appeared with Thought Bubble 1, pictures of cake and presents appeared with Thought Bubble 2) and putting those parts together (Thought Bubble 3) to make an inference. This idea was reinforced throughout the modules--whenever scaffolding and feedback on an inference was required, we illustrated TeLCI's thought process by using thought bubbles to demonstrate how connecting two ideas could lead to a third idea—i.e., an inference.

Teachers thought that students should have a second chance to answer questions. This issue relates to Instructional Design Principle 6, that students should have sufficient opportunities to respond. We agreed that a second chance to respond to questions would support student learning; thus, we modified the feedback so that, the first time a student answered a question incorrectly, TeLCI would explain why the answer was incorrect, and provide the students with another chance to answer the question.

Students found the scaffolding confusing and didn't always need it. This issue relates to Instructional Design Principle 7, that the instructional components should be carefully sequenced. We realized that presenting the scaffolding *before* each question was potentially confusing, and was not always needed, taking up valuable time. Thus, we changed the scaffolding so that, instead of presenting it before asking the questions, students could access it *after* hearing each question by (a) pressing a "Help" button before answering the question or (b) answering the question incorrectly and receiving scaffolding before their second attempt at answering the question.

The modules took slightly longer than ideal (25-30 min instead of 15-20 min). This issue related to our criterion to ensure that TeLCI was feasible to implement in classroom contexts. By altering scaffolding and feedback approaches as described above, we streamlined

the TeLCI modules so that they could be completed in 15-20 min.

Students had some difficulty navigating within the TeLCI application. This issue related to our criterion to ensure that students could use the software application independently in the classroom. To address this problem, we worked with the software team to improve intuitiveness and functionality of buttons used to navigate through the modules.

At the end of Year 1 and into Year 2, after identifying key problems and their solutions, we revised the existing four TeLCI modules and developed an additional 20 modules (for a total of 24; 12 fiction and 12 nonfiction). Figure 4 shows the revised sequence of questioning with scaffolding and feedback. A demo of the final version of the TeLCI components can be viewed here: https://innovation.umn.edu/tutoring-system/.

[Figure 4 about here]

Step 6: Conduct a small field trial to determine optimal timing of questioning. Following the development and refinement of the entire set of TeLCI modules, in Year 2 of the project we conducted a field trial to determine optimal timing of module questions (*online*--during viewing versus *offline*--after viewing). A detailed description of the field trial is beyond the scope of this paper and can be found elsewhere (Kendeou, McMaster, et al., 2018); here we provide a general overview of how it fit into our development process.

Setting and participants. This field trial was conducted in one elementary school located in a Midwestern metropolitan school district. Because TeLCI was designed as a Tier 2 intervention for students identified as at risk specifically for comprehension difficulties, we screened 107 students with parental consent (80 first- and 27 second-grade students), and identified 67 students as low- to low-average on the Understanding Spoken Paragraphs subtest of the Clinical Evaluation of Language Fundamentals-5th Edition (CELF-5; Wiig,

Semel, & Secord, 2003). To identify these students, we used cut-off criteria suggested by the

CELF-5 manual for identifying children at risk.

Measures and procedures. The 61 students completed pretest web-based video

assessments aligned to the intervention (i.e., *proximal* TeLCI assessments that involved

watching videos similar to those used in the TeLCI modules and answering inferential

questions either online or offline, but without scaffolding or feedback). Students then

engaged in three TeLCI modules per week for eight weeks. TeLCI modules were

implemented by classroom teachers, with technical support provided by the research team as

needed. The intervention was used during independent learning time and typically occurred

during early-morning hours of the school day. Teachers tended to rotate 3 to 5 students at a

time to work individually on a TeLCI module; each student took about 15-20 min to

complete a full module, with all students completing a full module within an hour.

Following the 8-week intervention, students were post-tested using the proximal

TeLCI assessment and the CELF-5 Understanding Spoken Paragraphs subtest. Findings

indicated that students in both conditions made significant gains over time on the TeLCI

proximal assessment, and that offline questioning ($d = 0.76$) was slightly more beneficial

than online questioning ($d = .58$) in improving inferencing skills. On the more distal

measure (CELF-5 subtest), findings indicated that students in both conditions made

significant gains over time. Similar to the proximal measures, student performance on the

CELF-5 showed a pattern whereby offline questioning ($d = 0.44$) was slightly more

beneficial than online questioning ($d = .30$) in improving language comprehension skills.

Students indicated they preferred offline questioning, reinforcing students' preferences

expressed in the Year 1 field test. These findings shed light on the optimal timing of

questioning to support comprehension in struggling young learners.

Recap of Design and Design Process

To summarize our process, we set out to develop a technology-based intervention system that was *theoretically*, *practically*, and *empirically* sound. To be theoretically sound, the intervention system was designed to meet the design principles described above and in Table 1. To be practically sound, we examined the feasibility and usability of the intervention system when implemented in authentic classroom settings. To be empirically sound, we examined a key component of the intervention—the timing of questioning—to ensure that the questioning approach was optimized to support inference-making.

Through a field test and field trial, we refined the system to ensure that children understood what inferences are and that they can be made across different media (Design Principle 1); that the timing of questions was optimized—by ultimately selecting the *offline* version (Principle 4); that they had sufficient opportunities to respond (Principle 6); and that the sequencing of questioning, scaffolding, and feedback supported their inference-making (Principle 7). Further, we revised the system to optimize feasibility/usability in terms of duration and navigability. With the application revised and refined to meet our design principles and criteria, we turned to developing the transfer component.

<p align="center">Development of Transfer Lessons</p>

In line with Instructional Design Principle 1, the purpose of the TeLCI transfer modules was to facilitate students' transfer of inference making from a non-reading context to a reading context--specifically, engaging in children's literature. Whereas strong evidence supports that inferencing transfers across media (Kendeou, 2015; Kendeou, Bohn-Gettler, White, & van den Broek, 2008), many students are more likely to transfer learned skills if

they are provided explicit teaching for transfer (Barnett & Ceci, 2002; Stokes & Baer, 1977).

Thus, an additional design principle for the transfer component of our intervention system is

that *transfer of inference-making from watching videos to engaging in read-alouds must*

be explicitly taught. Such instruction would include some common stimuli with the TeLCI

learning modules (e.g., the presence of the TeLCI character, the definition of inferencing,

presentation of vocabulary, questioning with scaffolding and feedback) but would be

delivered as a small-group read-aloud activity led by a teacher. The transfer lessons were

designed to be delivered at the end of each week that students completed three video-based

modules, which were grouped together based on an overarching theme (e.g., survival).

First, we identified an initial set of age-appropriate books that fit the goal-action-

outcome framework that guided video selection. We specifically selected books that aligned

with themes from the video-based learning modules that students would be completing in

the same week as the transfer lesson. As we did for the videos, we identified three key

vocabulary words that were central to the main ideas of each text. We wrote child-friendly

definitions and sample sentences that were relevant to each story.

Next, using procedures similar to those in the development of the video modules,

identified places in the text where inferences were needed to construct a coherent

representation of the text. We then wrote five inferential questions for each text and scripted

scaffolded prompts and feedback for incorrect responses. Unlike the questions for the video

modules, these questions were open-ended. In this way, students not only would have the

opportunity to make inferences receptively, by selecting from several constrained answers,

but also to make them expressively, by generating and articulating their own responses. For

each question, we included a note for the teacher to prompt students to share their answers to

questions with a partner before calling on one student to share his or her answer with the

group, to ensure that all students had opportunities to respond.

Scaffolded prompts include two steps: (1) the teacher repeated a relevant part of the

text as a reminder of what happened in the text followed by another opportunity to answer the

inferential question and (2) if the student still did not respond correctly, the teacher provided a

correct answer with an explanation of why it was correct. As with the video module feedback,

teacher feedback was tied to the type of response selected.

This process was completed for four transfer texts so that the transfer lessons could be

field tested and the process revised as needed before developing additional texts (our plan was

to develop a total of eight texts--four fiction and four nonfiction). We also wrote brief scripts

for teachers to introduce the transfer text, review what an inference is with additional

examples, and relate the transfer text to the theme of the video modules that students

completed that week. We created "Flat TeLCI"--a large picture of TeLCI on a poster board

mounted on a foam backing with an easel stand (see Figure 5). We also created three

laminated thought bubbles that the teacher could attach to TeLCI to demonstrate making an

inference during the transfer lessons. This way, the teacher could make direct connections to

the TeLCI learning modules, using common stimuli to facilitate transfer (Stokes & Baer,

1977).

[Figure 5 about here]

To field-test the transfer modules, graduate research assistants worked with a subset of

the TeLCI Year 2 field trial participants. We pulled participants into small groups of 3 to 5

students, and delivered the lessons in 20-min sessions. While one research assistant delivered

the lessons, another took field notes to inform revisions to the lessons. Through this process,

we identified several issues to address, including: (a) stories that were inappropriate due to too

many characters (and were thus difficult for students to follow), (b) parts of the transfer

lessons that were unwieldy or confusing, and (c) logistical aspects of how the modules were

packaged that were a detriment to implementation. Specifically, we discovered that our

question-scaffolding-feedback loop was long and difficult to follow. We attempted to

streamline this process making the scaffolding and feedback more clear and concise, and by

developing a flowchart that would help teachers decide what kind of feedback to provide

based on students' answers. See Figure 6 for the final version of this flowchart.

[Figure 6 about here]

After further revision, we invited several of our teacher partners to a focus group. In

this focus group, we demonstrated a transfer lesson, and asked teachers to review the

materials that we had developed. They provided further input into a lesson format that would

be usable and feasible to follow, as well as into the optimal group size (4-5 students) and how

they envisioned the lessons working in the context of the classroom (generally, they agreed

that the lessons could be delivered during daily guided reading time, when students typically

rotate through small groups and other center or independent work time).

Randomized Control Trial

The last step in our development process is to evaluate the promise of the intervention

system to improve student outcomes. At the time of the writing of this paper, we are in Year 3 of

the TeLCI project. In this final study year, we used a similar screening process as in the Year 2

field trial to identify students at-risk for language comprehension difficulties. These students ($n \sim$

60) were assigned randomly to the final version of TeLCI or to a business-as-usual control

group. TeLCI students completed eight weeks of learning modules and transfer lessons. All

students completed pre- and posttests of language and reading comprehension skills, using both proximal (TeLCI assessment) and distal (standardized language and reading comprehension assessment) measures. We will examine whether students' decoding, vocabulary, and executive functioning skills moderate their response to the TeLCI intervention. Evidence from this trial will be used to determine the need for further development and refinement of the intervention system.

Refining Our Process to Develop ELCII (Or, Lessons Learned and Applied)

As mentioned earlier, our aim is to develop an intervention system in the context of MTSS, with TeLCI targeting students identified as at-risk for language comprehension difficulties in Grades 1 and 2 (Tier 2), and ELCII reaching all students receiving core instruction in kindergarten (Tier 1). We have focused on TeLCI to illustrate our development process, which has played a significant role in ELCII development. Like TeLCI, ELCII has a strong theoretical and empirical foundation that builds upon prior findings of cognitive, developmental, intervention and assessment work. Successes and lessons learned from TeLCI development were adopted, refined, and applied to ELCII development. Two key lessons include the following.

Strengthen the Collaborative Relationship with School Partners

No amount of theory or empirical data will ensure success if classroom teachers are not fully invested in the intervention concept and implement it with fidelity. One example of strengthening the role of our teacher partners includes the selection of texts for transfer modules. During TeLCI development, we learned that appropriate text selection is a complex process, and so we enlisted the help of our teacher partners and asked them--as classroom experts--to select age appropriate texts that aligned with their standards and curriculum. These text selections serve as the basis for ELCII transfer module development.

Quality Control is Essential

Developing an intervention system to be delivered by a software application is a complex and multi-faceted endeavor, with many team-members contributing to many moving parts that must be checked, and double-checked, and checked again! Ultimately, we adopted a two-pronged approach to TeLCI quality control. First, we validated the accuracy of the content to ensure that all intervention material was accurately transferred from the research and development phase into the software application. Members of our research team reviewed the content of modules in written form and documented any inaccuracies. Then, we validated technical features to ensure that the intervention was working as intended in a user-friendly way.

Once all of the assessments and modules were validated for content and revisions completed, the TeLCI intervention was delivered to our software engineers who integrated it with the software application and tested basic functionality. The final quality control included our research staff viewing the intervention as a student would and troubleshooting any instances or issues that posed potential problems. Finally, we worked with our school partners to validate that the intervention was also functional "on site" (e.g., connecting on their Wi-Fi system). This quality control process ensured that the intervention performed as we had anticipated and alleviated any major problems during the field trial. These quality control systems are proving invaluable in our continued development and refinement of both TeLCI and ELCII.

<div align="center">

Conclusions

</div>

Inferencing is foundational to language and reading comprehension, and many children can benefit from early intervention aimed at developing this critical skill. Given that inferencing is a *general* skill that transfers across media, there is no need to wait for children to acquire fluent decoding skills before providing inferencing instruction. We propose that technology can

be leveraged to support students' inference making in non-reading contexts, and that their

learning in non-reading contexts can then transfer to reading contexts through explicit

instruction. Thus, we are developing the ELCII and TeLCI intervention systems, by building on

a strong theoretical foundation and providing both empirical and practical supports. Our hope is

that this system proves to be usable, feasible, and effective in promoting inferencing skills in

young children, and ultimately leads to improved reading comprehension.

References

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn?: A taxonomy for far transfer. *Psychological Bulletin, 128*, 612-637.

Butler, A. C., & Roediger, H. L. (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology, 19*, 514-527. doi:10.1080/09541440701326097

Cain, K., Oakhill, J. V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference making ability, and the relation to knowledge. *Memory & Cognition, 29*, 850-859. doi:10.3758/BF03196414

Catts, H. W., Compton, D., Tomblin, J. B., & Bridges, M. S. (2012). Prevalence and nature of late-emerging poor readers. *Journal of Educational Psychology*, *104*, 166-181. doi:10.1037/a0025323

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*, 354–80. http://doi.org/10.1037/0033-2909.132.3.354

Choi, S., & Clark, R. E. (2006). Cognitive and affective benefits of an animated pedagogical agent for learning English as a second language. *Journal of Educational Computing Research, 34*, 441-466. doi: 10.2190/A064-U776-4208-N145

Cohn, N. (2018). In defense of a "grammar" in the visual language of comics. *Journal of Pragmatics, 127,* 1-19.

Cohn, N., & Bender, P. (2017) Drawing the line between constituent structure and coherence relations in visual narratives. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*, 289-301. doi:10.1037/xlm0000290

Domagk, S. (2010). Do pedagogical agents facilitate learner motivation and learning outcomes?.

     Journal of Media Psychology, 22, 84-97. doi: 10.1027/1864-1105/a000011

Elleman, A. M. (2017). Examining the impact of inference instruction on the literal and

     inferential comprehension of skilled and less skilled readers: A meta-analytic

     review. *Journal of Educational Psychology, 109*, 761-781. doi:10.1037/edu0000180

Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Hillsdale, NJ:

     Erlbaum.

Goldman, S. R. (2004). Cognitive aspects of constructing meaning through and across multiple

     texts. In N. Shuart-Faris & D. Bloome (Eds.), *Uses of intertextuality in classroom and*

     educational research (pp. 317-351). Charlotte, NC: Information Age Publishing.

Golke, S., Dörfler, T., & Artelt, C. (2015). The impact of elaborated feedback on text

     comprehension within a computer-based assessment. *Learning and Instruction, 39*, 123-

     136. doi:10.1016/j.learninstruc.2015.05.009

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*,

     81-112. doi:10.3102/003465430298487

Heidig, S., & Clarebout, G. (2011). Do pedagogical agents make a difference to student

     motivation and learning?. *Educational Research Review, 6*, 27-54.

     doi:10.1016/j.edurev.2010.07.004

Helder, A., Van Leijenhorst, L., & van den Broek, P. (2016). Coherence monitoring by good and

     poor comprehenders in elementary school: Comparing offline and online measures.

     Learning and Individual Differences, 48, 17-23. doi: 10.1016/j.lindif.2016.02.008

Kendeou, P., Bohn-Gettler, C., White, M. J., & van den Broek, P. (2008). Children's inference

    generation across different media. *Journal of Research in Reading, 31*, 259-272.

    doi:10.1111/j.1467-9817.2008.00370.x

Kendeou, P., McMaster, K. L., Butterfuss, R., Kim, J., Bresina, B., & Wagner, K. (in press). The

    inferential language comprehension (*iLC*) framework. Invited paper for *Topics in

    Cognitive Science*.

Kendeou, P., McMaster, K. L., & Christ, T. J. (2016). Reading comprehension: Core components

    and processes. *Policy Insights from the Behavioral and Brain Sciences*.

    doi:10.1177/2372732215624707

Kendeou, P., McMaster, K. L., & TeLCI Team (2018). Year 2 TeLCI Pilot Study. *Manuscript in

    preparation*.

Kendeou, P., van den Broek, P., White, M. J., & Lynch, J. S. (2009). Predicting reading

    comprehension in early elementary school: The independent contributions of oral

    language and decoding skills. *Journal of Educational Psychology, 101*(4), 765-778.

    doi:10.1037/a0015956

Kim, Y.-S. G. (2016). Direct and mediated effects of language and cognitive skills on

    comprehension or oral narrative texts (listening comprehension) for children. J*ournal of

    Experimental Child Psychology, 141*, 101-120. DOI: 10.1016/j.jecp.2015.08.003

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York, NY: Cambridge

    University Press.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A

    historical review, a meta-analysis, and a preliminary feedback intervention theory.

    Psychological Bulletin, 119, 254-284. doi:10.1037/0033-2909.119.2.254

Krippendorff, K. H. (2003). *Content analysis: An introduction to its methodology*. Thousand

Oaks, CA: Sage Publications.

Kulik, J. A., & Kulik, C. C. (1988). Timing of feedback and verbal learning. *Review of

Educational Research, 58*, 79-97. doi:10.3102/00346543058001079

Lorch, E. P., & Sanchez, R. P. (1997). Children's memory for televised events. In P.W. van den

Broek, P.J. Bauer, & T. Bourg (Eds.), *Developmental spans in event comprehension and

representation: Bridging fictional and actual events* (pp. 271–291). Hillsdale, NJ:

Erlbaum.

Lynch, J. S., & van den Broek, P. (2007). Understanding the glue of narrative structure:

Children's on- and off-line inferences about characters' goals. *Cognitive Development,

22*, 323-340. doi: 10.1016/j.cogdev.2007.02.002

Lynch, J. S., van den Broek, P., Kremer, K. E., Kendeou, P., White, M. J., & Lorch, E. (2008).

The development of narrative comprehension and its relation to other early reading skills.

Reading Psychology, 29, 327–365. doi:10.1080/02702710802165416

Magliano, J. P., Larson, A. M., Higgs, K., & Loschky, L. C. (2016). The relative roles of

visuospatial and linguistic working memory systems in generating inferences during

visual narrative comprehension. *Memory & Cognition, 44*, 207-219.

Magliano, J. P., Loschky, L. C., Clinton, J., & Larson, A. M. (2013). Is reading the same as

viewing? An exploration of the similarities and differences between processing text- and

visually based narratives. In B. Miller, L. Cutting, & P. McCardle (Eds.), *Unraveling the

behavioral, neurobiological, & genetic components of reading comprehension* (pp. 78-

90). Baltimore, MD: Brookes Publishing Co.

McMaster, K. L., & Espin, C. A. (2017). Reading comprehension instruction and intervention: Promoting inference making. D. Compton, R. Partial, & K. Cain (Eds.). *Theories of reading development.* Amsterdam, Netherlands: John Benjamin's Publishing.

McMaster, K. L., van den Broek, P., Espin, C. A., Pinto, V., Janda, B., Lam, E. A., … van Boekel, M. (2015). Developing a reading comprehension intervention: Translating cognitive theory to education practice. *Contemporary Educational Psychology, 40,* 28-40. doi:10.1016/i.cedpsych.2014.04.001

McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes*, *22*, 247–288.

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, *14*, 1–43.

Oakhill, J. (1984). Inferential and memory skills in children's comprehension of stories. *British Journal of Educational Psychology*, *54*, 31-39.

Recht, D. R., & Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology*, *80*, 16–20.

Rohrer, D., & Taylor, K. (2007). The shuffling of mathematics problems improves learning. *Instructional Science*, *35*, 481–498. http://doi.org/10.1007/s11251-007-9015-8

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 119-144. doi:10.1007/BF00117714

Stokes, T. G., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*. doi: 10.1901/jaba.1977.10-349

Van den Broek, P., Helder, A., & Van Leijenhorst, L. (2013). Sensitivity to structural centrality:

    Developmental and individual differences in reading comprehension skills. In M. A.

    Britt, S. R. Goldman, & J.-F. Rouet (Eds.), *Reading: From words to multiple texts* (pp.

    132-146). New York, NY: Routledge, Taylor & Francis Group.

Van den Broek, P., Lorch, E.P., & Thurlow, R. (1996). Children's and adults' memory for

    television stories: The role of causal factors, story-grammar categories, and hierarchical

    level. *Child Development, 67*, 3010–3028. doi:10.1111/j.1467-8624.1996.tb01900.x

Wiig, E. H., Semel, E. M., & Secord, W. (2003). *CELF 5: Clinical Evaluation of Language

    Fundamentals*. Pearson/PsychCorp.

Zacks, J. M., Speer, N. K., & Reynolds, J. R. (2009). Segmentation in reading and film

    comprehension. *Journal of Experimental Psychology: General, 138*, 307-327.

    doi:10.1037/a0015305

Table 1

Instructional Design Principles Applied to Developing and Refining the Intervention

Instructional Design Principle	Principle Applied to the Intervention	Revisions Based on Field Test & Field Trial
1. Inference skills should transfer across media.	We developed modules to teach what an inference is, and to provide opportunities for children to make inferences while watching videos. Later, we developed transfer lessons to apply inference-making in a read-aloud context.	We refined how the concept of inferencing was introduced and reinforced throughout the video modules.
2. Making inferences requires background knowledge.	We identified key vocabulary and concepts needed to make appropriate inferences while watching the videos, to be presented prior to watching each show.	Students and teachers expressed satisfaction with this component and indicated that it was helpful, so no substantive revisions were made.
3. Questioning should improve inference making.	We identified key points in each video where inferences were needed to understand the show, and developed "why" questions that would prompt children's generation of the appropriate inferences.	Students and teachers expressed satisfaction with this component, and students appeared to make growth over time when they received the TeLCI modules, so no substantive revisions were made.
4. The timing of questioning should be optimized.	We developed and compared the effects of *online* questions (asked during the videos) and *offline* questions (asked after the video) to determine the optimal timing.	Patterns in data from the field trial suggested a slight benefit of offline questioning, and students expressed a preference for this approach, so the final version of modules and transfer lessons have offline questions.
5. Scaffolding and feedback should allow for the systematic training of inferences.	We identified key events in the video that should be activated and integrated to make appropriate inferences, and used these to develop scaffolding cues to support inference making. We also developed explicit feedback to indicate whether a child's answer was correct or incorrect, along with an explanation for the correct response.	Students indicated that scaffolding and feedback were helpful to their learning. However, see revisions to Instructional Design Principle #7 to improve the sequence of questioning, scaffolding, and feedback.
6. Inference training should provide multiple opportunities to respond.	We developed 5 questions for each video to allow students to have multiple opportunities to respond.	Teachers indicated that students should have a second chance to respond to a question when their first response was incorrect. Thus, in our revision we added a second chance whenever a question was answered incorrectly.
7. Inference training should be carefully sequenced.	We initially developed each module to follow a sequence in which TeLCI (a) introduced vocabulary, (b) showed a video segment (online) or entire video (offline), (c) provided scaffolding, (d) asked an inferential question, and (e) provided feedback.	Some children did not need the scaffolding and the scaffolding made the modules unnecessarily long. Thus, we revised the sequence so that scaffolding followed a question only if a student's response was incorrect.

Table 2

Demographic Data for Field Test Participants

	Grade 1 (*n*=71)	Grade 2 (*n*=20)	Total (*n*=91)
Gender	53.5% Female	50% Female	52.7% Female
Ethnicity	54.9% White, 21.1% Hispanic 18.3% African American 2.8% American Indian 2.8% Asian	5% White, 10% Hispanic 85% African American 0% American Indian 0% Asian	44% White, 18.7% Hispanic 33% African American 2.2% American Indian 2.2% Asian
Age	M=7.03 (SD=.34)	M=8.02 (SD=.46)	M=7.27 (SD=.56)
Free/Reduced Lunch Status	46.5%	90%	56%
Special Education Status	8.5%	10%	8.8%
Home Language	79% English 11% Spanish 2% Gujarati 1% Arabic 1% Tibetan 1% Oromo 1% Amharic	79% English 12% Spanish 2% Gujarati 1% Arabic 1% Tibetan 1% Oromo 1% Amharic	79% English 12% Spanish 2% Gujarati 1% Arabic 1% Tibetan 1% Oromo 1% Amharic

Note. Due to rounding, percentages may not add to 100. Demographic data could not be obtained for 1 of the first-grade students and 4 of the second-grade students in the study.

Table 3

Categories Derived from Field Test Observations

| Categories	Definition	Notes
App features & functions • Sequence of activities • Buttons • Instruction/ practice • Loading • Navigation	Observation related to student's success in using the app to operate and navigate the module	Most issues relate to timing and function of buttons or navigation problems. Some issues relate more to students' confusion about sequencing and scripting of activities and instructions.
Attention/engagement • Disengaged • Distracted • Engaged • Semi-engaged • Short span	Observation related to whether student seemed engaged with TeLCI	Of these observations, 40% noted students' disengagement/ distraction, 42% noted students' engaged behavior, and the rest 18% noted semi-engagement, or impatience/short attention span.
Device	Observations related to students' interaction or difficulties with hardware (tablets, mouse, headphones, trackpad)	Some students required extra support in learning how to use the devices.
Online vs. offline	Students expressed a preference for one or the other	Offline: 56% preferred Online: 44% preferred
Questions • Comments about specific questions • Level of difficulty • Number • Selection options • Overall	Students commented on aspects of the questions.	Most commented about the level of difficulty (but it was mixed as to whether they were too easy or too hard). Several commented that there were too many. Some thought the questions were "just right" and others didn't like them.
Overall evaluation	Observers noted students' general comments about whether they liked the TeLCI activities.	The vast majority were positive. Some students thought it was too long or wouldn't want to do it again.

Table 4

Summary of Student Survey Responses

Question (*How helpful was … to your learning with TeLCI?*)	Very Helpful	Kind of Helpful	Not very Helpful
Vocabulary	63.79%	27.59%	8.62%
Scaffolding	44.83%	41.38%	13.79%
Offline questions	52.63%	24.56%	22.81%
Online questions	55.17%	18.97%	25.86%
Feedback	50.00%	27.59%	22.41%
Help moving through app	55.17%	27.59%	17.24%
Help using buttons	60.34%	20.69%	18.97%
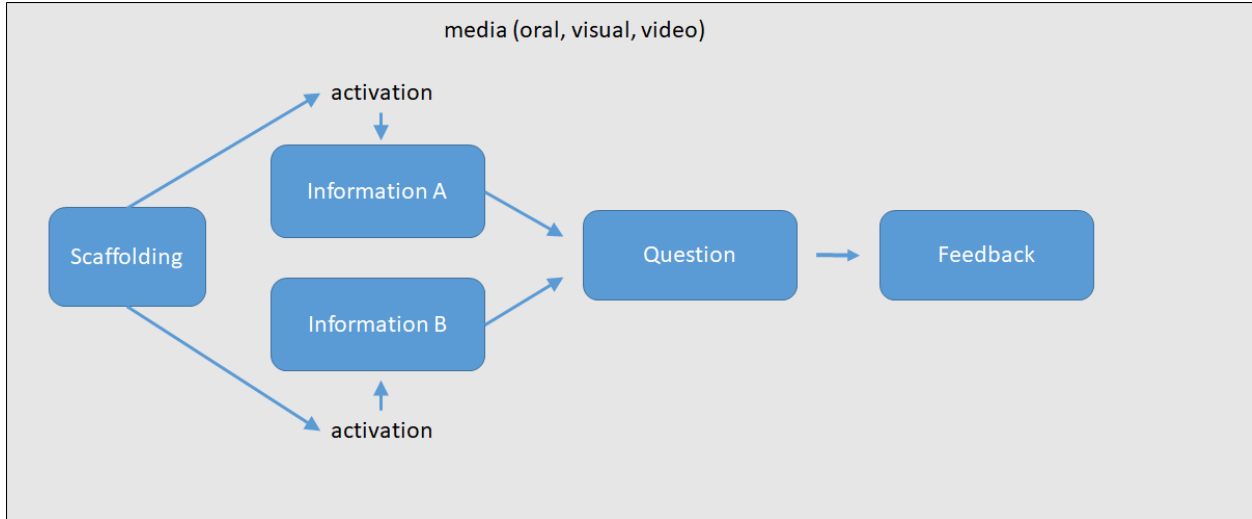
Panel A
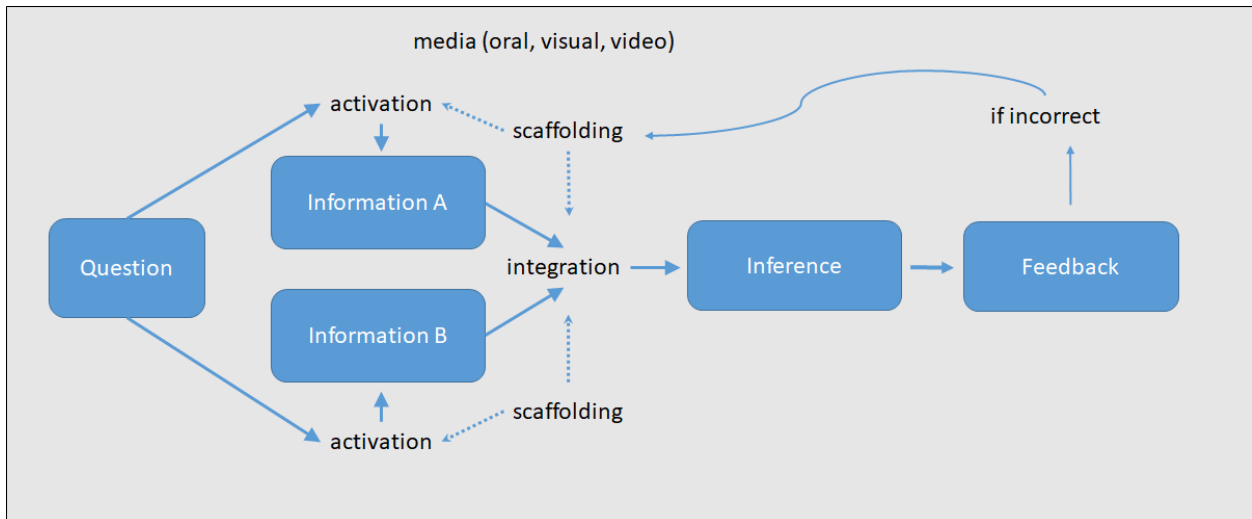
Panel B

Figure 1. Initial (Panel A) and final (Panel B) instructional sequence for questioning, scaffolding, and feedback.

1. **Scaffolding 1:** *Did you notice when Blinky Bill fell off the balloon and was hanging from the rope? Click "yes" or, if you're not sure, click "rewind" to watch that part again.*
   a. **If student clicks "yes:"** *I noticed that, too! Blinky WAS hanging from a rope outside the balloon!*
   b. **If student clicks "rewind:"** *Okay, let's rewatch that part.* (scaffolding video 1 plays) *So Blinky WAS hanging from a rope outside the balloon!*
2. **Scaffolding 2:** *Did you notice when the airplane went by and a passenger took a picture of Blinky Bill? Click "yes" or, if you're not sure, click "rewind" to watch that part again.*
   a. **If student clicks "yes:"** *I noticed that, too! A passenger on the airplane DID take Blinky's picture!*
   b. **If student clicks "rewind:"** *Okay, let's rewatch that part.* (scaffolding video 2 plays) *So a passenger on the airplane DID take Blinky's picture!*
3. **Model Making an Inference:** *So, we know Blinky was hanging from a rope outside the balloon, and we know a passenger on the airplane took Blinky's picture. Now, I'll ask a question that helps us connect these two ideas to make an inference.*
4. **Inferential question:** *The Circus Brothers found out where Blinky and his friends were going because they…*
   a. *saw Blinky's picture in the newspaper* (correct)
   b. *saw Blinky Bill flying in the sky* (incorrect)
   c. *talked to Blinky Bill on the phone* (incorrect)
   d. *thought Blinky was from Antarctica* (incorrect)
5. **Feedback:**
   a. **Feedback for option "a:"** *I agree! We can make an inference that the Circus Brothers knew where Blinky Bill and his friends were going because they saw Blinky's picture in the newspaper. Nice job making that inference!*
   b. **Feedback for option "b:"** *Hmmm, no, that can't be right. The Circus Brothers never actually saw Blinky Bill flying through the sky, so that can't be how they knew where he and his friends were going. The Circus Brothers knew where Blinky Bill and his friends were going because they saw Blinky's picture in the newspaper.*

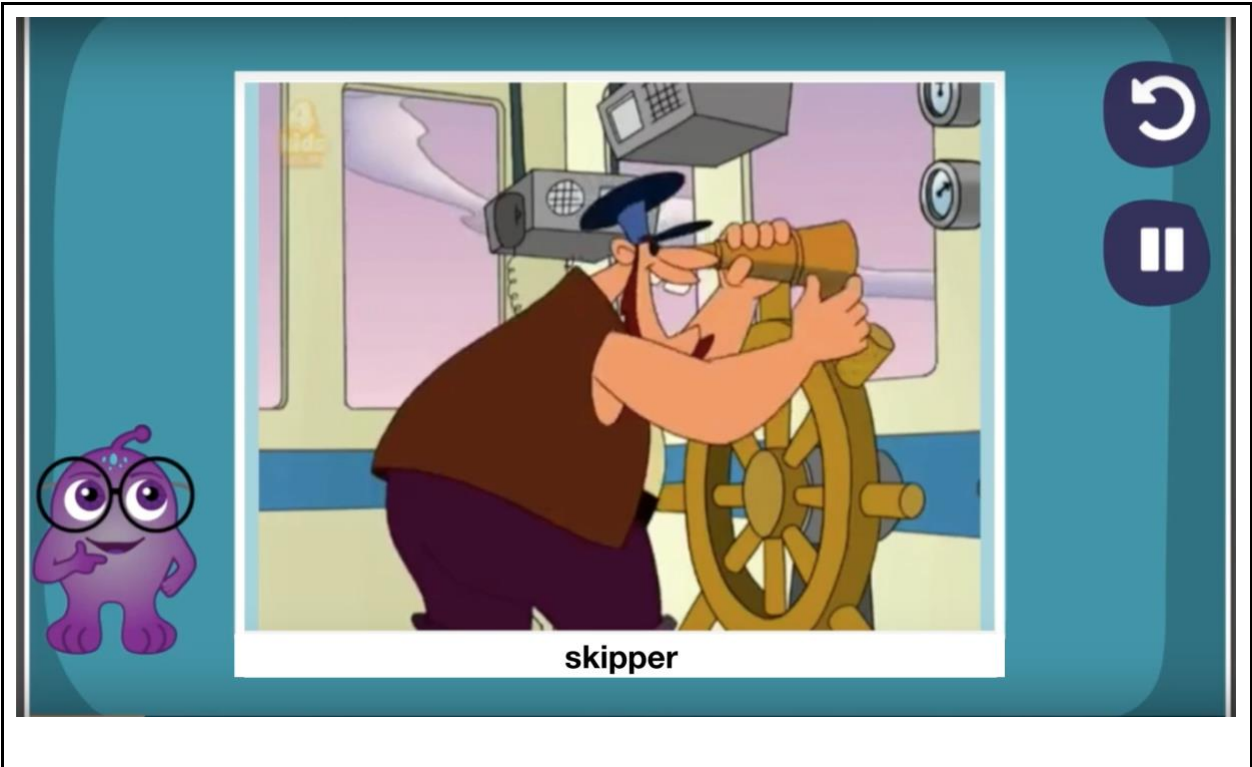*Figure 2.* Original TeLCI questioning sequence.

Figure 3. TeLCI vocabulary instruction.

1. **Inferential question:** *The Circus Brothers found out where Blinky and his friends were going because they…*
   a. *saw Blinky's picture in the newspaper* (correct)
   b. *saw Blinky Bill flying in the sky* (incorrect)
   c. *talked to Blinky Bill on the phone* (incorrect)
   d. *thought Blinky was from Antarctica* (incorrect)
   Choose your answer or, if you want help, click on the 'help' button.
2. **If student clicks the help button or answers the question incorrectly on the first attempt:**
   a. *Click 'rewind' to rewatch an important part that will help us make and inference.* **Scaffolding 1 (re-watch video clip):** *Blinky was hanging from a rope outside the balloon!*
   b. *Click 'rewind' for another important part of the show.* **Scaffolding 2 (re-watch video clip):** *So we know Blinky Bill was hanging from a rope outside the balloon* **(show Thought Bubble 1)** *and we know that a passenger on the airplane took Blinky's picture* **(show Thought Bubble 2).**
   c. **Show Thought Bubble 3:** *Now, let's connect these two ideas to make an inference and answer the question.*
3. **Second attempt at the inferential question after receiving scaffolding if it was needed:**
   The Circus Brothers found out where Blinky and his friends were going because they…
   a. *saw Blinky's picture in the newspaper* (correct)
   b. *saw Blinky Bill flying in the sky* (incorrect)
   c. *talked to Blinky Bill on the phone* (incorrect)
   d. *thought Blinky was from Antarctica* (incorrect)
6. **Feedback:**
   c. **Feedback for option "a:"** *I agree! We can make an inference that the Circus Brothers knew where Blinky Bill and his friends were going because they saw Blinky's picture in the newspaper. This is important because the circus brothers are trying to catch Blinky, and the picture in the newspaper told them where Blinky and his friends were! You made a good inference!*
   d. **Feedback for option "b"--after first try:** *Hmmm, that doesn't seem right. The Circus Brothers were not on the airplane so they never actually saw Blinky Bill flying through the sky. That can't be how they knew where Blinky and his friends were going.* **Play scaffolding.**
   e. **Feedback for option "b"--after second try:** *Hmmm, that doesn't seem right. The Circus Brothers were not on the airplane so they never actually saw Blinky Bill flying through the sky. That can't be how they knew where Blinky and his friends were going. The Circus Brothers knew where Blinky Bill and his friends were going because they saw Blinky's picture in the newspaper. This is important because the circus brothers are trying to catch Blinky, and the picture in the newspaper told them where Blinky and his friends were!*

Figure 4. Revised TeLCI questioning sequence.

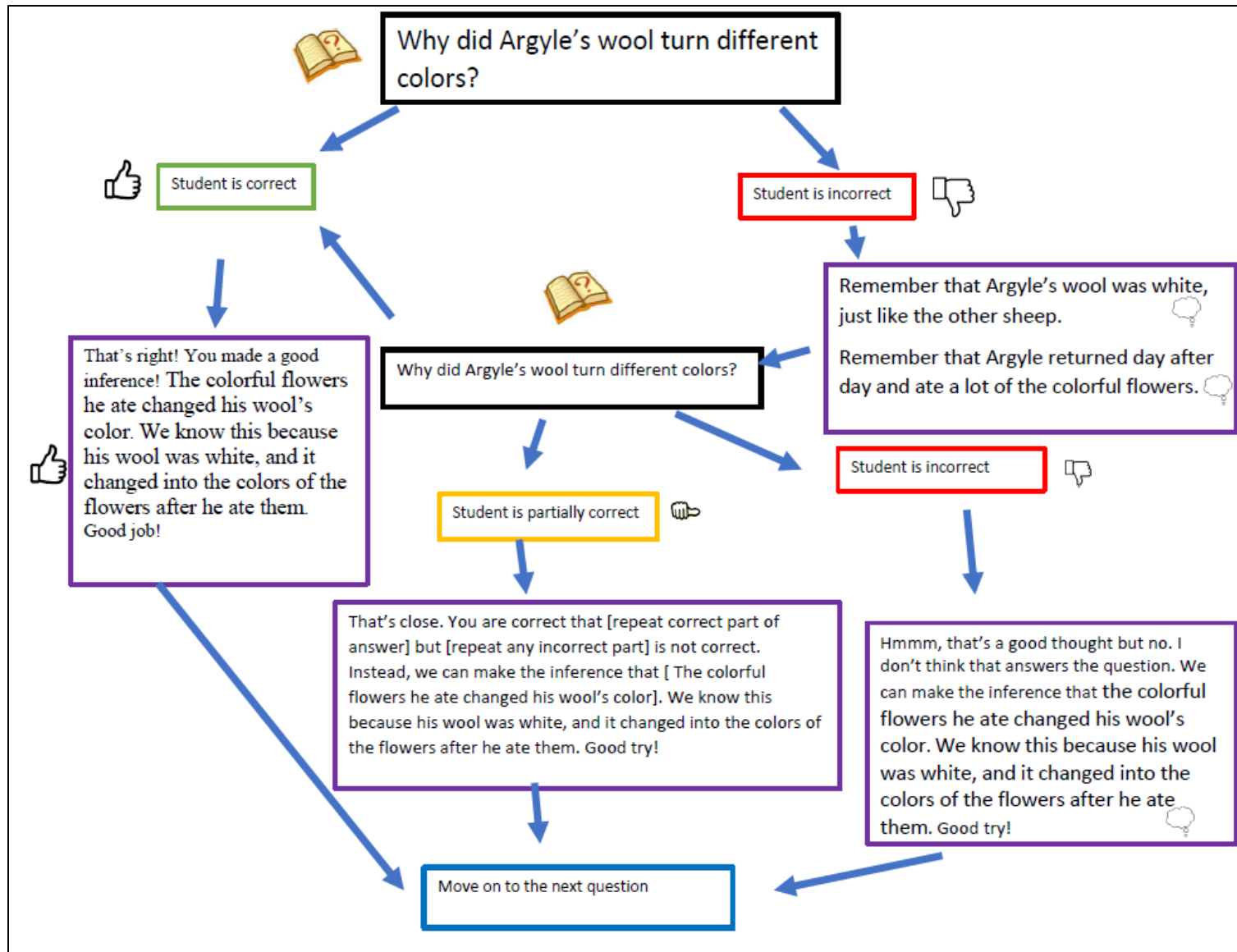*Figure 5*. TeLCI Transfer Materials.

Figure 6. Transfer flow chart.