Improving Reading Comprehension in Spanish using iSTART-E: A Pilot Study

Kathryn S. McCarthy  
*Georgia State University, USA*

Christian Soto  
*Universidad de Concepción, Chile*

Antonio Gutiérrez de Blume  
*Georgia Southern University, USA*

Diego Palma  
*Universidad de Concepción, Chile*

Jordan González  
*Universidad de Concepción, Chile*

Danielle McNamara  
*Arizona State University, USA*
Abstract

iSTART-E is a web-based intelligent tutor developed for Spanish-speaking students to improve their reading comprehension through self-explanation strategy training. This study examined the effects of a blended comprehension strategy intervention on students’ reading comprehension skill. Chilean high school students (n = 22) completed nine iSTART-E sessions and a face-to-face classroom lesson that included integrative video and additional examples of self-explanation. Survey data indicated that students thought that iSTART-E was useful and that they had improved their reading skills. Critically, this was supported by objective assessment -- students’ standardized reading comprehension test (Lectum) performance improved from pre-training to post-training. These findings demonstrate promise for the use of iSTART-E as a computer-supported learning environment for Spanish readers.

Keywords: Reading comprehension, reading strategies, artificial intelligence, tutoring system, educational technology.
Improving Reading Comprehension in Spanish using iSTART-E

Reading is a fundamental skill. As such, improved literacy vastly expands individuals’ possibilities of progress in school and working life (OECD, 2013). Thus, the improvement of students’ literacy skills (e.g., the ability to learn from text and clearly communicate ideas) should be a central goal of schooling around the world (Paris & Hamilton, 2009). Unfortunately, international assessments indicate that many students struggle with reading comprehension (NAEP, 2015; San Martín, Jara, Preiss, Claro, & Farina, 2012; Shea & Ceprano, 2017). Of particular interest to our research team was the development of reading comprehension support for Chilean students. Although Chile performed as well or better than other Latin America countries on the PISA (an international assessment of 15-year-olds performance in reading, mathematics, and science), scores in 2018 were 35 points lower than the Organisation for Economic Co-operation and Development (OECD) average. Statistics show that 28.4% in 2015 and 21% in 2018 of Chilean students did not achieve basic level of reading comprehension (OECD, 2016, 2019).

These numbers demonstrate the need for scalable reading comprehension interventions for Spanish-speaking Chilean students. Computer-supported language learning environments offer the means to support students’ literacy skills at scale. Despite the growing numbers of computer-based systems, there are relatively few systems that support higher-order reading comprehension processes for Spanish speakers (cf. Vidal-Abarca et al., 2014; Veliz & Osorio, 2001; Ponce, López, & Mayer, 2012). Thus, we introduce the Interactive Strategy Training for Active Reading and Thinking en Español, of iSTART-E. The purpose of the present study was to explore the effects of iSTART-E in a classroom setting – both in terms of student perceptions and objective learning outcomes. We first outline the theoretical motivations for reading comprehension strategy instruction and the foundations of computer-supported self-explanation training. We conclude with initial evidence for the efficacy of iSTART-E in an authentic classroom setting.

Improving Reading Comprehension through Self-Explanation

Reading and learning from text require a variety of complex cognitive processes. In addition to lower level processes, such as decoding and sentence parsing, reading also involves higher order comprehension processes (McNamara & Magliano, 2009). Theories of discourse comprehension (e.g., Graesser, Singer & Trabasso, 1994; Kintsch, 1988) suggest that readers need to infer connections between information from across different parts of the text and to integrate information from the text with information in long-term memory (i.e., prior knowledge) in order to construct a coherent mental model of a text. Encouraging students to generate inferences and to draw upon information from prior knowledge can improve comprehension.

One of the most effective means of improving comprehension is through self-explaining, or explaining the text to oneself during reading (see Bisra, Liu, Nesbit, Salimi & Winne, 2018). Indeed, skilled readers spontaneously produce self-explanations as they read (Cote, Goldman & Saul, 1998). Prompting less-skilled readers to self-explain encourages them to monitor their own understanding and to connect ideas from across the text with information that they already know. Making these inferences between parts of the text and between text and prior knowledge improves comprehension and learning (Allen, McNamara & Mc Crusden, 2015).

The modest benefits of prompting students to self-explain can be enhanced by training students to generate high quality self-explanations. More specifically, students can leverage
comprehension strategies that encourage the students to more actively engage with the text and to make inferences that support deeper comprehension. Relevant to the current work are five strategies (described below) that have been shown to improve reading comprehension: comprehension monitoring, paraphrasing, predicting, bridging, and elaboration.

Comprehension monitoring refers to being aware of one’s own understanding (McNamara, 2007; NRP, 2000; Palincsar & Brown, 1984, 1986; Paris et al., 1986). Skilled readers (and indeed, skilled learners more generally) tend to monitor their comprehension more so than less skilled readers (Côté, Goldman, & Saul, 1998; Goldman et al., 2012). Less skilled readers also tend to be less accurate in their metacognitive judgments because they rely on superficial cues (fluency, ability to recall) rather than meaningful comprehension (Theide, Griffin, & Wiley, 2010). By increasing the rate that students’ attend to their understanding and helping them to focus on what it means to “understand”, they are more likely to work to repair comprehension problems using other reading strategies (Chi & Bassok, 1989; Connor et al., in press; McNamara, 2004; Oakhill et al., 2005). In this way, comprehension monitoring is a critical aspect of successful reading comprehension. Paraphrasing, or putting the text into your own words, externalizes the reader’s understanding and encourages the reader to activate relevant prior knowledge (Best et al., 2005). Paraphrasing is an important first step toward deeper processing. Prediction, or thinking about what might be coming next in the text, encourages readers to think ahead and think about the text as a whole. The final two strategies bridging and elaboration are related in that they require students to make inferences (Singer, 1998). Bridging links ideas from across sentences to help “fill in the blanks” left open by the text (e.g., McNamara et al., 1996). In order to successfully comprehend a text, the reader must generate bridging inferences to build a coherent mental model that connects the separate ideas across the text (Kintsch, 1988). Elaboration is the process of making inferences that link what is in the text or sentence to related knowledge. For example, when reading this sentence, “The blood vessels that extend across the heart and supply it with blood are called the coronary arteries.”, a reader might connect this information to their knowledge of the circulatory system and how it provides oxygen to the entire body. The reader might also use general knowledge or logic to infer that a blockage in these arteries might have a variety of implications for a person’s health. Encouraging readers to use logic and common sense helps them to understand that it is possible to make sense of the text, and go beyond the text, without knowing a lot about the topic (McNamara, 2004).

In the Self-Explanation Reading Training (SERT; McNamara, 2004, 2017) intervention, students are taught to self-explain using a variety of comprehension strategies (monitoring, paraphrasing, inferencing) that have been shown to improve reading comprehension. Critically, students also engage in extended practice with feedback to develop their self-explanation skills. That is, after initial instruction, students read other texts and practice using the different comprehension strategies in their self-explanations. Peers and teachers then help them to generate higher quality explanations that involve inferencing. Consistent with theories of skill acquisition (e.g., Ericcson, Krampe, & Tesch-Romer, 1993), these cycles of deliberate practice and feedback are essential for learning to self-explain and, in turn, to increase reading comprehension skill.

Computer-Supported Strategy Training
Although SERT is a powerful intervention, it is best suited to small group instruction so that an instructor can provide tailored feedback as students’ practice using the comprehension strategies
in their self-explanations. Moving training to a computer-based environment affords the opportunity to deliver instruction at scale.

Generally speaking, computer-based tutoring systems have been shown to be effective, demonstrating gains similar to one-on-one or small group tutoring (Kulik & Fletcher, 2016; Ma, Adesope, Nesbit, & Liu, 2014; Steenbergen-Hu & Cooper, 2014; VanLehn, 2011). A limitation to early computer-based tutoring was that it was limited to well-structured domains (e.g., math, science) in which the system can use close-ended assessment items to build the learner model. However, advances in natural language processing have allowed researchers and designers to develop algorithms that can quickly assess and provide feedback for open-ended responses (Graesser & McNamara, 2012; Passoneau et al., 2018). The Interactive Strategy Training for Active Reading and Thinking (iSTART; McNamara, Levinstein, & Boothum, 2004; see also McCarthy et al., in press; Snow, Jackson, Jacovina, & McNamara, 2016) is part of a growing body of tutoring systems specific reading and writing (Crossley & McNamara, 2016; Strobl et al., 2019). These systems sometimes use close-ended items, but also employ analysis of word and semantic-based features of students’ written products (e.g., written verbal protocols, short-answer questions, summaries, essays) to deliver individualized feedback and support. A recent meta-analysis, specific to intelligent tutoring systems for reading, indicates that these computer-based environments yield similar learning gains as other systems designed for other domains (Xu et al., 2019).

Interactive Strategy Training for Active Reading and Thinking: iSTART

The Interactive Strategy Training for Active Reading and Thinking (iSTART; McNamara, Levinstein, & Boothum, 2004; see also Snow, Jackson, Jacovina, & McNamara, 2016) is a game-based intelligent tutoring system built on the foundations of SERT. iSTART provides self-explanation training and practice games for five comprehension strategies: comprehension monitoring, paraphrasing, prediction, bridging, and elaboration. We describe the features and architecture of iSTART in the context of iSTART-E below, but, in brief, iSTART uses video lessons, guided demonstration, and game-based practice to help students learn to write high quality self-explanations. In order to provide automated evaluation of these open-ended self-explanations, iSTART relies on state-of-the-art natural language processing algorithms to identify the strategies used in each self-explanation. The algorithm provides a score from (0-3) as well as feedback on how the learner can improve the quality of their self-explanation. The ability to deliver self-explanation reading training in a computer-supported learning environment means that students can receive one-on-one instruction, support, and feedback in ways that would otherwise be impossible in the classroom. Critically, the system also includes game-based practice as a means of promoting students’ motivation, including their interest, engagement, enjoyment, and self-efficacy in the task(s) (Jackson & McNamara, 2013). At its simplest, motivation matters because students who refuse to use the system will not be able to reap any of its benefits and the longer students are willing to use the system, the more they have the opportunity to engage in extended practice. Further, motivation influences the types of processes engaged during learning activities (Garris, Ahlers, & Driskell, 2002).

iSTART has been shown to improve reading comprehension for English speakers across middle, high school, and college students (Snow, Jacovina, Jackson & McNamara, 2016). However, an obvious limitation for iSTART is that it is limited to students who read and learn in predominantly English language texts. This motivated us to develop iSTART en Español to support literacy for Spanish speaking students.
**iSTART-E: Architecture**

The design of iSTART-E is modeled on iSTART. In iSTART-E, students first watch video lessons that introduce self-explanation and the five reading comprehension strategies; comprehension monitoring, paraphrasing, prediction, bridging, and elaboration (descriptions and examples of each of the strategies appear in Table 1).

Figure 1 shows the lesson video menu and Figure 2 shows a screenshot of one of the comprehension strategy lesson videos. After each lesson, students take a checkpoint quiz to ensure their understanding of each strategy.

Table 1.

*Description of each Self-Explanation Strategy and Examples From Spanish Self Explanations*

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Description</th>
<th>Example Self-Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comprehension</td>
<td>maintaining awareness of one’s own understanding.</td>
<td>“Entiendo que la contracción de los músculos provoca que el pecho se expanda, durante una respiración profunda”.</td>
</tr>
<tr>
<td>Monitoring</td>
<td>Comprehension monitoring allows readers to repair</td>
<td>(I understand that the contraction of the muscles causes the chest to expand, during a deep breath).</td>
</tr>
<tr>
<td></td>
<td>comprehension problems using other reading strategies.</td>
<td></td>
</tr>
<tr>
<td>Paraphrase</td>
<td>restating information directly from the text in</td>
<td>“El pecho se expande cuando se contraen los músculos”.</td>
</tr>
<tr>
<td></td>
<td>different words.</td>
<td>(The chest expands when the muscles contract)</td>
</tr>
<tr>
<td>Prediction</td>
<td>thinking about what might be coming next in the text</td>
<td>“Creo que a continuación el texto comentará cómo el oxígeno provee de energía al cuerpo”.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(I think that the next part of the text will comment on how oxygen provides energy to the body).</td>
</tr>
<tr>
<td>Bridge</td>
<td>linking ideas between the current sentence and other</td>
<td>“El aire se mueve hacia el interior de los pulmones cuando el pecho se expande y se mueven las costillas hacia arriba y hacia afuera”.</td>
</tr>
<tr>
<td></td>
<td>sentences in the text</td>
<td>(The air moves to the lungs when the chest expands and the ribs move up and out).</td>
</tr>
<tr>
<td>Elaboration</td>
<td>making inferences that link what is in the text</td>
<td>“Cuando corro, tengo que tomar respiraciones profundas, que permitan que el aire entre en mis pulmones y el torrente sanguíneo entregue energía a mi cuerpo”.</td>
</tr>
<tr>
<td></td>
<td>related prior knowledge</td>
<td>(When I run, I have to take deep breaths, which allows the air to enter to my lungs and the bloodstream give energy to my body).</td>
</tr>
</tbody>
</table>
Figure 1. Screenshot of iSTART-E’s Menu

Figure 2. Screenshot of iSTART-E’s lesson phase
Once students complete the lesson videos, they are directed to a practice environment, “Práctica Dirigida” (Directed Practice; Figure 3). In Práctica Dirigida, students read a text and are prompted to write self-explanations for target sentences (indicated in Figure 3 in bold). The self-explanation is scored by the algorithm and a pedagogical agent provides a score from “pobre/poor” (0) to “genial/great” (3) as well as formative, actionable feedback to revise and improve the self-explanation. Students are then given the opportunity to revise their self-explanation using this feedback.

![Figure 3. Screenshot of Práctica Dirigida](image)

After one round (one full text) of Práctica Dirigida, students are given access to an open practice environment. Students can re-watch the videos and go through additional rounds of Práctica Dirigida, but they may also engage in game-based practice. Game-based practice is included in iSTART-E to keep students engaged and motivated (see Jackson & McNamara, 2013). This includes generative games, in which students write their own self-explanations, and identification games, in which students view example explanations and identify the strategy that was used. In the generative game Conquista del Mapa (Map Conquest), students write self-explanations to earn flags. Higher self-explanation scores earn more flags to place on the board. Students try to conquer more squares on the board than their CPU opponents. At the end of each game, students’ points are converted to iFichas, the in-system currency. iFichas can be used to purchase plays of identification games. In identification games (Partido de Estrategia, Constructor de Puentes, Estallido del Globo, and Escape del Calbozo; Figures 4 and 5), each game has its own mechanics for play, but students advance in each game by correctly identifying the strategy in an example self-explanation.
**Figure 4.** Screenshot of the game Escape del Calabozo

**Figure 5.** Screenshot of Estallido de Globos

**iSTART-E: Development**

The transformation of iSTART to iSTART-E was an immense undertaking. While the programming for the interface could be borrowed from iSTART, the lessons and texts needed to be translated to Spanish. A native Spanish speaking member of the team led the translation efforts to ensure that the video lessons and texts were not simply translated word-by-word, but
rather that the Spanish version reflected the same meanings and nuances in ways that were culturally-relevant and appropriate (Soto et al., 2015).

The other major issue was the need to develop a self-explanation scoring algorithm. The iSTART algorithm is derived from English language tools or tools that rely on English language corpora (e.g., Coh-Metrix [McNamara et al., 2014], Latent Semantic Analysis [Landauer et al., 1998]) to derive indices. Thus, we needed to produce a Spanish-specific algorithm that could detect when students produce lower-level self-explanations such as paraphrases, simple comprehension monitoring statements (scores of 0 or 1) as compared to when they engage in inferencing (2 or 3).

To mirror this scoring procedure, we created a genetic-based (evolutionary) algorithm that analyzes word and discourse-based features (see Dascalu, Jacovina, Soto, Allen, Dai, Guerrero, & McNamara, 2017, Table 1) of Spanish language to match human ratings of self-explanation quality. This algorithm yielded 69.5% exact accuracy and 94.1% adjacent accuracy with human raters (Dascalu et al., 2017). Thus, iSTART-E is able to automatically score and provide feedback on Spanish language self-explanations in a similar manner to one-on-one instruction.

The Current Study
Although we took great care to found iSTART-E in the same theoretically-based and empirically-validated design of iSTART, it is inappropriate to assume that Spanish language reading comprehension training will operate identically to English iSTART. Design frameworks (e.g., ADDIE, DBIR) suggest conducting iterative, small-scale evaluations, prior to large-scale implementations (Fishman et al., 2013; see also Stone et al., 2018 for design research specific to intelligent tutoring systems). With this in mind, we elected to conduct a small-scale study to explore implementation and effectiveness of iSTART-E. These data can be used to demonstrate the promise of iSTART-E training as well as to provide feedback for how to improve the system as it grows.

Thus, the current study was an initial examination of the efficacy of the lessons and practice in iSTART-E in an authentic Spanish-speaking classroom. We have reported preliminary pretraining/postraining data from this study in McCarthy et al. (2018). The present study provides a deeper description of the study and a richer analysis of students’ experiences with iSTART-E in a Chilean classroom. The intervention involved high school students engaging in multiple rounds of instruction and practice with iSTART-E as well as a classroom-based face-to-face training session that supported the lessons taught in the iSTART-E system. We used a multi-pronged approach that employed both objective and subjective evaluations in order to assess multiple aspects of system efficacy. This included post-training survey to examine students’ experiences and perceptions of iSTART-E. It also included proximal evaluations such as a knowledge test to explore students’ explicit knowledge of the strategies and analysis of students’ self-explanation over the course of training. The study also employed a standardized reading comprehension assessment, Lectum as a more distal outcome. In addition, we collected data related to students’ reading motivations to explore how individual differences might impact system efficacy.

Method

Participants
Forty-five students from a 1° Medio (first year of high school) course at a school in Nacimiento, Chile were recruited for this study. We omitted those students who 1) were absent on the day of pretest, 2) absent on the day of posttest, and/or 3) missed one of the training days. Thus, we retained the 22 students who completed the pretest and posttest, as well as all iSTART-E sessions. The ages of these 22 students ranged from 14 to 17 (M_{age} = 15, SD = 0.75).

Measures

Pre-training and Post-training Reading Comprehension Assessment

Theories of reading comprehension (e.g., Graesser, Singer, & Trabasso, 1994; Kintsch, 1988) posit that comprehension emerges from the construction of a coherent mental model or mental representation of a text. This mental model includes not only information explicit in the text, but also inferences needed to connect information from across different parts of the text as well from the text to information in prior knowledge. Researchers often refer to explicit memory of information as “shallow comprehension” whereas “deep” comprehension emerges when one has a robust and interconnected mental model that includes many inferences. Previous work has shown that SERT and iSTART are particularly effective for improving students deep comprehension (McCarthy et al., 2018; McNamara, 2004, 2017). Thus, it was critical to include a reading assessment that evaluates both shallow and deep aspects of comprehension.

Lectum is a reading comprehension test for the Chilean school system. Lectum assesses various cognitive components of the reading process (Riffo, Véliz, Castro, Reyes, Figueroa, Salazar, & Herrera, 2011; Soto, Gutiérrez de Blume, Contreras, & Carrasco, 2019). There are two equivalent paper-based forms of Lectum (A and B). Each test consists of five texts belonging to different discursive genres. Readers complete 40 multiple-choice items. The test takes approximately one hour (60 minutes) to complete.

Lectum includes text-based items (questions for which the answer can be found explicitly in the text) designed to evaluate shallow comprehension as well as inference items (questions for which the reader must infer information through meaningful comprehension of the text) to assess deeper comprehension. Lectum Form B was administered prior to the intervention (pretest) and Form A was administered after training (posttest).

Motivations for Reading Questionnaire

The Motivations for Reading Questionnaire (MRQ; Wigfield & Guthrie, 1997) captures differences in students reading motivations. These motivations relate to students’ reading frequency (Wigfield & Guthrie, 1997) and are related to improvements in students’ reading comprehension skill (Guthrie et al., 2007). The MRQ consists of 53 Likert-scale items (1-4) and yields 11 subscores (sample items appear in italics):

1) Reading Efficacy (I am a good reader)
2) Reading Challenge (I like it when the questions in books make me think)
3) Reading Curiosity (I read to learn new information about topics that interest me)
4) Reading Involvement (I feel like I make friends with people in good books)
5) Importance of Reading (It is very important to me to be a good reader)
6) Work Avoidance (I don’t like reading something when the words are too difficult)
7) Competition (I like to finish my reading before other students)
8) Recognition for Reading (My parents often tell me what a good job I am doing in reading)
9) Reading for Grades (I read to improve my grades)
10) Social Reasons for Reading (*I talk to my friends about what I am reading*)
11) Compliance (*I do as little schoolwork as possible in reading*)

**Perception Survey**
At the end of each 45 to 55 minutes session, students completed a brief Likert-style (1-5) survey that evaluated their *experience* and *perceptions* of the system and that day’s training, as well as *transfer perception* items that asked students to report if they had applied the strategies taught and trained by iSTART-E in other texts or contexts (Table 6). These items were designed to quickly evaluate various aspects of student motivation.

**Strategy Knowledge Test**
Immediately following the classroom-based, face-to-face lesson related to the iSTART-E comprehension strategies, students completed a strategy knowledge test via Google Forms. Students were first asked to identify which one of the five strategies was used in an example. They were then asked to provide a definition of that strategy. Students were given an example of each of the five strategies. Thus, they could receive a 0 (incorrectly identified) or 1 (correctly identified) for application of a strategy and a 0 (incorrect definition) or 1 (correct definition) for the definition of the strategy – for each of the five strategies.

**Procedure**
The iSTART-E intervention consisted of 11 total sessions. Session 1 included only the pre-training Lectum assessment and the MRQ. The final session, Session 11, included only the post-training Lectum assessment. In the nine training sessions, students engaged with iSTART-E, starting the video lessons first and then progressing to the open practice environment. After each session they completed a perception survey.

In the middle of the intervention, students engaged in a group learning activity. They watched and analyzed a video developed by the research team (Figure 6). This video summarized the five comprehension strategies in iSTART-E: monitoring, paraphrasing, prediction, bridging, and elaboration.

![Figure 6. Screenshot of The New Integrative video (Session 4)](image)
After the video, students clarified doubts, reviewed examples, and proposed various situations in which they could use each of these strategies. They then complete the strategy knowledge test.

**Results**

The effects of iSTART-E were explored through multiple outcome measures. Objective measures included analysis of the quality of the self-explanations produced during training, assessment of students’ explicit knowledge of the comprehension strategies, and a standardized reading assessment. Subjective measures included survey data related to the motivations, experiences, and sense of system efficacy. These data collectively suggest that iSTART-E supported students’ enjoyment and sense of learning, as well as their actual learning outcomes.

**Self-Explanation Quality**

We used log data (data collected by the iSTART system during use) to inspect the nature of the self-explanations that students generated during practice. Early in the training, students’ self-explanations are generally quite weak. For example, Table 2 shows student data from their first iSTART-E session. The table includes an excerpt from the text (target sentence indicated in bold) and the student’s self-explanation. The score (1) was generated by the iSTART-E algorithm (Dascalu et al., 2017). This self-explanation stays close-to-text rather than connecting information from prior text (bridging) or integrating information with prior knowledge (elaboration).

Table 2.  
**Example text and self-explanation from early in iSTART-E training**

<table>
<thead>
<tr>
<th>Text</th>
<th>Student’s Self-Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Por ejemplo, la radiación gamma es capaz de causar mutaciones.</strong></td>
<td>La radiación gamma, también puede causar mutaciones. (iSTART Score: 1)</td>
</tr>
<tr>
<td>[For example, gamma radiation is capable of causing mutations.]</td>
<td>[Gamma radiation can also cause mutations].</td>
</tr>
</tbody>
</table>

Later in training, students’ self-explanations tend to be longer and more complex, characterized by the use of connectors that establish different relationships between ideas. Table 3 includes a self-explanation generated by the same student in a later session. Although the target sentence remains relatively similar, the student is making clear connection to previous parts of the text.

Table 3.  
**Example text and self-explanation from late in iSTART-E training**
La temperatura mide el promedio de energía de una cierta cantidad de moléculas, y está relacionada con el promedio de velocidad con la que se mueven las moléculas. Cuanto mayor sea la temperatura, más rápida (en promedio) es la velocidad en la que se mueven las moléculas. Se llama punto de fusión a la temperatura en la cual el agua cambia de estado sólido a uno líquido.

[The temperature measures the average energy of a certain number of molecules, and this is related to the average speed with which the molecules move. As the temperature is higher, faster (on average) is the speed in which the molecules move. It’s called a melting point to the temperature in which water changes from a solid to a liquid state.]

Similar gains were observed across a majority of the students. These findings suggest that instruction and practice with the five comprehension strategies helped students to generate higher quality self-explanations.

**Overall Reading Skill Gains**

While the log data demonstrates gains on the proximal skill of self-explanation, it was also important to assess the more distal, but target outcome of improved reading comprehension skill. Thus, we examined overall pretest to posttest changes in *Lectum* reading comprehension test scores. A t-test indicated that students’ reading comprehension skill significantly improved from pretest ($M = 37.32, SD = 22.42$) to posttest ($M = 75.36, SD = 25.87$), $t(21) = 5.21, p < .001$, Cohen’s $d = 1.57$.

**Reading Motivation**

As shown in Table 4, students in this sample were neutral to slightly below neutral on most of the motivation scales. They appear to value reading, but view themselves as just okay, rather than good or great readers. There were no significant correlations between MRQ subscores and reading comprehension gains from pretest to posttest. However, it is of note that some reading motivation scales (e.g., Efficacy, Importance) were more strongly positively correlated with gains, whereas other aspects of motivation (e.g., Involvement) were negative correlated.

### Table 4.

**MRQ sub-score relation with reading gains**

<table>
<thead>
<tr>
<th>MRQ Sub-score</th>
<th>$M$</th>
<th>$SD$</th>
<th>$r$ with Lectum Gains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading Efficacy</td>
<td>2.73</td>
<td>0.77</td>
<td>0.33</td>
</tr>
<tr>
<td>Reading Challenge</td>
<td>2.86</td>
<td>0.64</td>
<td>-0.06</td>
</tr>
<tr>
<td>Reading Curiosity</td>
<td>2.91</td>
<td>0.61</td>
<td>0.17</td>
</tr>
</tbody>
</table>
Strategy Knowledge
Strategy knowledge tests administered after the classroom-based session asked students to identify and apply each of the five comprehension strategies. For each strategy, each student’s response was scored as correct (1) or incorrect (0). Table 5 shows the percentage of students who correctly answered each question. In general, the majority of students were able to correctly identify and apply each of the strategies. The only exception to this was defining elaboration, in which fewer than half of students were able to correctly define the elaboration strategy. These findings suggest that students who have completed iSTART-E training have a strong grasp on effective reading comprehension strategies.

Table 5.
Table. Percentage of student who correctly answered questions about each of the comprehension strategies

<table>
<thead>
<tr>
<th>Knowledge Test</th>
<th>Monitoring</th>
<th>Paraphrasing</th>
<th>Prediction</th>
<th>Bridging</th>
<th>Elaboration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applying Strategy</td>
<td>90.91%</td>
<td>63.64%</td>
<td>77.27%</td>
<td>77.27%</td>
<td>59.09%</td>
</tr>
<tr>
<td>Defining Strategy</td>
<td>68.18%</td>
<td>72.73%</td>
<td>90.91%</td>
<td>63.64%</td>
<td>40.91%</td>
</tr>
</tbody>
</table>

Attitudes
Students’ attitudes and perceptions of their experiences with iSTART-E were generally positive (Table 6). They found the system to be useful and effective and they reported relatively high motivation and effort in the iSTART-E activities. However, it is of note that students reported less favorable evaluations of the text assessment (scores) and the games. Thus, we intend to examine these aspects more closely to better understand how to make improvements.

Table 6.
Table. Average iSTART-E Experience and Attitude Ratings (1-5) across All Sessions

<table>
<thead>
<tr>
<th>Question</th>
<th>M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technologies are useful for reading comprehension learning</td>
<td>4.09 (0.87)</td>
</tr>
<tr>
<td>iSTART-E was effective</td>
<td>4.23 (0.75)</td>
</tr>
<tr>
<td>How motivated were you when doing the activities?</td>
<td>3.45 (0.96)</td>
</tr>
<tr>
<td>How much effort did you put to progress in the activities?</td>
<td>3.68 (0.72)</td>
</tr>
<tr>
<td>Was the iSTART-E experience good?</td>
<td>3.68 (0.78)</td>
</tr>
<tr>
<td>How was the design of the tool?</td>
<td>3.45 (0.96)</td>
</tr>
<tr>
<td>How was the text assessment in the tool?</td>
<td>2.45 (0.60)</td>
</tr>
</tbody>
</table>
How did you like the games? 2.45 (0.74)

Students were also asked about how using iSTART-E had affected their learning. Students reported feeling like their reading comprehension skill had improved; that they were able to apply the comprehension strategies outside of iSTART-E training; and that they found these strategies useful (Table 7).

Table 7. Progress perception questionnaire

<table>
<thead>
<tr>
<th>Question</th>
<th>M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improvement perception</td>
<td>4.06 (0.85)</td>
</tr>
<tr>
<td>Did you apply the strategies learned so far?</td>
<td>3.91 (0.75)</td>
</tr>
<tr>
<td>Do you think learning strategies are useful?</td>
<td>4.18 (0.73)</td>
</tr>
</tbody>
</table>

To summarize, students’ attitudes indicated generally favorable response to iSTART-E both in terms of motivation and in terms of perceived efficacy. Students’ perceptions of improvement were reflected in analysis of their self-explanations and in the LECTUM data, in which students showed significant gains in reading comprehension skill from pre-training to post-training. Students also demonstrated explicit knowledge of the comprehension strategies they had learned. Collectively, these findings provide evidence that iSTART-E may serve as an effective computer-supported learning environment for Spanish language reading comprehension training.

Discussion

The Interactive Strategy Training for Active Reading and Thinking en Español (iSTART-E) is a Spanish language computer-supported learning system for reading comprehension strategy training. This system was developed in part to help meet the demand for scalable Spanish-language reading comprehension interventions. iSTART-E provides video lessons on five different comprehension strategies that students can employ as they self-explain the texts that they read. iSTART-E uses a natural language processing-based evaluation system (Dascalu et al., 2017) to provide immediate, individualized feedback to help students increase the quality of their self-explanations and, in turn, their reading comprehension skill. The system also includes game-based practice to increase student motivation.

Although iSTART-E is an adaptation of iSTART (an English-language system) that has demonstrated consistent positive effects on reading comprehension, it was important to collect empirical data to establish the promise of the iSTART-E system on its own. The current study reflects an initial small-scale, classroom-based efficacy test. Subjective and objective measures collected during our classroom study demonstrate the promise of iSTART-E training for developing Spanish-speaking students’ reading comprehension skill. Students’ attitudes and perceptions of their experiences indicated that students reacted positively to most aspects of
iSTART-E training and, perhaps more importantly, found the training to be effective and valuable for their learning. Despite a small sample size ($N_{iSTART-E} = 22$) and only 10-12 hour of training, there was evidence of improvement on self-explanation quality over the course of training and robust gains in standardized reading comprehension performance ($d = 1.57$).

**Future Directions**

The intent of this study was to examine the promise of iSTART-E on this small sample, prior to larger implementations. Given the positive results, we aim to conduct further empirical research to examine the scalability and generalizability of these effects. In addition to larger sample sizes across a variety of settings, it will also be important to conduct experimental comparisons to explore how iSTART-E training compares to various control conditions, such as no training or business-as-usual instruction. Notably, iSTART-E, like its sister system, is intended to supplement, rather than replace, reading comprehension strategy instruction. As such, we will also conduct research to examine how system-only iSTART-E training compares to different types of blended instruction that includes both iSTART-E computer-supported training and classroom-based instruction and peer interaction. These future directions are important, both in terms of system development and refinement and in better understanding reading comprehension intervention for Spanish-language readers. Consistent with aspects of design-based research, we intend to continue to test the system across a variety of settings and learners.

**Conclusions**

The rapid growth of natural language processing tools has afforded researchers and designers the opportunity to provide computer-based reading comprehension instruction for a wide audience. iSTART-E addresses a need to provide targeted literacy instruction in Spanish-speaking classrooms. Our findings suggest that computer-based instruction and practice, in concert with minimal face-to-face lessons, can serve to provide enjoyable and efficacious reading comprehension instruction. By leveraging educational technologies, like iSTART-E, researchers and educators can support high-quality student learning at scale.

**REFERENCES**


OECD (2016). Informe de resultados PISA 2016, Chile.


