Student-Generated Questions in Literacy Education and Assessment

Lois Maplethorpe1, Hyunah Kim1, Melissa R. Hunte1, Megan Vincett1, and Eunice Eunhee Jang1

Abstract
This study investigated the extent to which students’ questioning ability is associated with their literacy abilities, attitudes, perceived text understanding, and interest in the text they read. We further examined these relationships by the type of text they read to generate questions. Fifth- and sixth-grade students (N = 89) were asked to generate three questions after reading two different types of text. The students also completed reading comprehension and writing tests, as well as a questionnaire about their attitude toward literacy, perceived text understanding, and interest in the text. A hierarchical regression analysis showed that the quality of student-generated questions was predicted by reading comprehension ability, a positive attitude toward writing, and perceived level of understanding of the text, with strong effects related to text genre. We explore the implications of these findings on current pedagogy and assessment practices in literacy education and suggest areas for further research.

Keywords
student-generated questions, reading comprehension, metacognition, literacy assessment, literacy processes

Introduction
Question generation before, during, and after reading can be a valuable activity as it taps into student discourse in “talking literacy” (Bugg & McDaniel, 2012; Davey & McBride, 1986; Taboada & Guthrie, 2006). There is an increasing interest in student-generated questioning (SGQ) as it can provoke students’ critical thinking, stimulate their curiosity and interest, and foster their navigation into various disciplines (Chin

1Ontario Institute for Studies in Education of the University of Toronto, Toronto, Canada

Corresponding Author:
Lois Maplethorpe, Department of Applied Psychology and Human Development, Ontario Institute for Studies in Education of the University of Toronto, 252 Bloor Street West, Toronto, ON, Canada M5S 1V6. Email: lois.maplethorpe@mail.utoronto.ca
& Brown, 2002; Moje, 2008; Shanahan & Shanahan, 2008). Through SGQ, students can explore novel ideas that they may not have considered.

SGQ can also provide formative information about students’ conceptual confusions, gaps, or misunderstandings (Chin & Brown, 2002). This allows for a more targeted approach to learning where teachers can realign lessons or adjust strategies to address discrepancies. As such, SGQ has the potential to capture diversity among students, providing more authentic and meaningful evidence of learning and understanding than traditional, assessor-directed methods (Duncan & Buskirk-Cohen, 2011; Guthrie & Cox, 2001). Approaches to teaching and assessment that involve SGQ help to generate active, engaged learners (Taboada & Guthrie, 2006).

However, typical classroom pedagogy and assessment practices have largely ignored the opportunity for students to initiate questions. Cazden (1988) characterized the traditional classroom discourse sequence of teacher initiation (I), student response (R), and teacher evaluation (E) as the IRE pattern, which illuminates the “asymmetry in the rights and obligations of teacher and students” (p. 54) over the right to speak in typical classrooms.

This asymmetrical discourse extends to assessment beyond classrooms, where students typically respond to questions whose correct answers are predetermined by assessors. When students respond to teacher-posed questions, they “read to satisfy the teachers’ purposes, not their own” (Singer & Donlan, 1982, p. 171). Cognitive discordance arises between generating questions and responding to externally generated questions (Chin & Brown, 2002; Taboada & Guthrie, 2006). Indeed, as one Swedish science student lamented, “the trouble with school…is that it provides uninteresting answers to questions we have never asked” (Osborne, 2006, p. 2). Equally troubling is that teachers’ questions often fail to activate students’ critical, epistemological, and reflective thinking beyond factual and procedural information (Becker, 2000; Stone, 2017). Instead, teachers tend to ask factual questions, which fall short of revealing the breadth and depth of understanding that students may actually have or desire to know (Jang, 2014; Resnick, 1987).

Students’ engagement in social contexts plays an important role in their literacy development (Guthrie & Wigfield, 2000); however, teacher/assessor-led evaluative discourse may impact “motivated literacy” (Moje, 2008; Unrau et al., 2014). In particular, such discourse is shown to have profound differences among students from different linguistic, cultural, and socioeconomic backgrounds (Au, 1980; Keehne et al., 2018), prompting a call for a shift from teacher- to student-led discourse (Cazden, 1988).

Despite relatively rich research on SGQ in the fields of science and mathematics education, there is a dearth of research on whether elementary school students’ SGQ is associated with literacy development and attitudes. The current study addressed this gap by examining the role of SGQ in elementary school students’ literacy learning. Specifically, we examined its associations with young students’ reading and writing abilities, their attitude toward reading and writing, their perceived understanding of the text, and their interest in passage topics. We further examined whether the predictive relationships differ by text genre, comparing results in narrative and expository texts.
Literature Review

Literacy Development

Literacy lays the foundation for individual growth, academic success, and career advances. Building on skills required to learn to read, school-aged children engage in increasingly dynamic and complex literacy activities beyond printed text materials (Cain et al., 2000; Organisation for Economic Co-operation and Development, 2019; RAND Reading Study Group, 2002). Individual readers are the agents of comprehension who bring abilities and prior experiences to the act of reading. Whereas word-reading skills play a critical role in reading achievement in kindergarten through Grade 3 (Torgesen & Hudson, 2006), starting in Grade 4, higher order comprehension processes, such as inference making and discourse-level integration, become central (Cain et al., 2000; RAND Reading Study Group, 2002).

Alongside more complex comprehension skills, metacognitive monitoring skills are considered the driving force behind developing later reading abilities (Koda, 2005). Metacognition involves the ability to self-regulate one’s own learning by setting goals, monitoring comprehension, executing repair strategies, and evaluating comprehension strategies. Research consistently supports the significant and positive role of metacognitive abilities in developing literacy (Cain et al., 2000).

Use of SGQ in Education

Students’ active questioning can be an important comprehension strategy that puts readers in an active, purposeful role when learning from diverse texts (Taboada & Guthrie, 2006). It may facilitate students’ sense of autonomy and control over learning as they become aware of what they can or cannot comprehend, how effective their reading strategies are, and how confident they are in their reading abilities (Taboada et al., 2012). Inquiry-based literacy activities can encourage meaning-focused and “socially just” learning for students of all reading abilities and diverse backgrounds (Learned, 2018, p. 190).

SGQ has been used in a variety of classroom contexts in science and mathematics. Researchers have found that question quality influences depth of knowledge construction and cognitive engagement (Chin & Brown, 2002; Feldt et al., 2002). Employing a case study approach within a Grade 8 science class, Chin and Brown (2002) analyzed student dialogue to investigate the relationship between student-generated questions (SGQs) and students’ approaches to learning. They found two types of SGQs emerged: factual information-seeking questions and deeper “wonderment” questions (Chin & Brown, 2002, p. 531). Although only 13% of the SGQs were the wonderment type, these questions triggered deeper-level thinking such as reflection, curiosity, puzzlement, skepticism, and speculation. The researchers concluded that SGQ can “play an important role in engaging the students’ minds more actively, engendering productive discussion, and leading to meaningful construction of knowledge” (p. 540).
SGQ has also been used in mathematics and science to improve students’ ability to formulate and solve complex problems. Gonzales (1996) presented preservice teachers with a math stimulus, such as a graph or chart, and asked them to generate five questions that could be answered directly from the stimulus or by extending knowledge beyond the graph or chart. Results indicated that the majority of SGQs were lower level queries, requiring a response directly related to the given prompt. However, a limited number of questions required more complex processes, such as modifying, extending, or adding ideas to the given information (Gonzales, 1996). Demirdogen and Cakmakci’s (2014) analysis of over 1,000 student questions to identify interest in chemistry, underlying motivations for asking the questions, and the cognitive complexity of the questions indicated a strong relationship between SGQs and student interest in learning.

Bates et al. (2014) and Jones (2019) applied SGQ to assessment by asking students to generate questions about course material prior to an exam/quiz. This activity had a positive effect on exam achievement for both high- and low-performing students. However, low-performing students appear to benefit the most from SGQ on subject comprehension measures (Bates et al., 2014). Jones (2019) suggested that the role-reversing process positions students as active agents in constructing and driving their own learning. This in turn supports more complex thinking since students evaluate, synthesize, and critique rather than simply memorizing or applying course material.

Another pedagogical use of SGQ that has received considerable attention is targeting reading comprehension (Bugg & McDaniel, 2012; Davey & McBride, 1986; Taboada & Guthrie, 2006). Increased attention to the text required for SGQ seems to engage more complex metacognitive processes, enhancing explicit and inferential comprehension (Davey & McBride, 1986). A metacognitive process related to reading comprehension is students’ perceived understanding of a text they have just read. Bugg and McDaniel (2012) found that SGQ enhanced high school students’ self-evaluation of text comprehension adequacy upon reading expository texts. Using a rating scale, students predicted how accurately they responded to post-passage comprehension questions. Students demonstrated higher prediction accuracy when generating conceptual questions than the control group, who only reread the passages (Bugg & McDaniel, 2012). The authors suggested that SGQs facilitated active engagement with the text, which allowed them to metacognitively monitor their comprehension and rectify any comprehension inadequacy on their own. In addition, Cohen (1983) found that Grade 3 language learners trained in SGQ showed significant improvement in literal-level comprehension test scores after responding to their own questions compared to the untreated control group. This suggests that further SGQ training among younger students can significantly improve their text comprehension skills.

SGQ is also used to improve students’ writing and is often utilized by proficient writers as an idea-generating strategy (Flower & Hayes, 1981). Research reveals that SGQ fosters increased text engagement, which allows for deeper, more insightful inferences and more meaningful, robust persuasive written arguments (Verlaan et al., 2014). Furthermore, the cognitive aspect of writing involves constructing links between prior and new knowledge (Oxford, 1990), evaluating claims (Holliday et al., 1994), and
integrating and applying meaning to real-world situations (Levin & Wagner, 2005). SGQs also comprise many of these cognitive, connective, generative, constructive, and integrative aspects, and given the overlap between writing processes and question generation, it is worth examining whether writing ability predicts the quality of questions students produce.

**Factors Associated With the Quality of SGQs**

Motivational beliefs, such as one’s attitude toward and interest in reading and writing, may be associated with the quality of SGQs. Students who enjoy reading and writing are intrinsically motivated and readily engage in the learning process, gaining ample practice (Guthrie & Cox, 2001). Students who are self-motivated to learn new things use more sophisticated, complex strategies. They metacognitively monitor their comprehension processes and are more likely to generate higher quality questions (Botsas & Padeliadu, 2003). Guthrie and Cox (2001) studied science classrooms that used a combination of student choice, hands-on learning, SGQ, and direct instruction in comprehension strategies to enhance motivation and engagement in reading. As students observed and explored their curiosities, they raised their own thought-provoking questions and enthusiastically sought out responses to these meaningful queries (Guthrie & Cox, 2001). While existing research provides some evidence that supports the positive influence of SGQ on students’ abilities and attitudes, there is little research that examines the relationship between various student characteristics and SGQ, especially for young students.

In addition, text genre is shown to have a moderating effect on SGQ quality, as each text structure may activate different thinking processes that call for different kinds of responses (Bugg & McDaniel, 2012; Kraal et al., 2018). For example, narrative texts call for the processing of thematic structures and knowledge-based elaborations, whereas expository texts incite more attention to text structure, technical vocabulary, and abstract concepts (Kraal et al., 2018; Roehling et al., 2017). More research is needed to understand how different text types prompt variations in SGQ quality among students and further moderate this relationship with learner variables.

**Evaluating the Quality of SGQs**

Previous research has used a binary rating scheme to evaluate SGQ quality based on literal versus inferential questions (Davey & McBride, 1986) and factual versus conceptual (Bugg & McDaniel, 2012). For example, Davey and McBride (1986) identified SGQs as either literal or inferential, where they were scored as correct if the response involved inferencing related to gathering the “gist” of the passage or integrating ideas from multiple sentences, and incorrect if the response to the question came directly from the text.

Other studies have used a more complex hierarchical scale, such as Bloom’s taxonomy (Bates et al., 2014; Taboada et al., 2012). The cognitive dimension of Bloom’s taxonomy (Bloom et al., 1956) comprises six hierarchical categories, from basic, concrete concepts to complex and abstract cognitive engagement. For example, Taboada
and Guthrie (2006) evaluated students’ ecological science questions using a six-level knowledge hierarchy where lower level questions were about isolated facts and higher level questions reflected more complex relationships and integration of concepts. Krathwohl’s (2002) expansion of Bloom’s work includes a bidimensional, hierarchical taxonomy that dichotomizes knowledge processes (Factual, Conceptual, Procedural, and Metacognitive) and cognitive processes (Remember, Understand, Apply, Analyze, Evaluate, and Create). This revised framework has been used by the Program for International Student Assessment (Organisation for Economic Co-operation and Development, 2019) and other large-scale assessments.

Besides Bloom’s taxonomy, there are various literacy-based alternative frameworks and approaches that can be used to typify SGQs. As Resnick (1987) posited, complex thinking and questioning can be generated through everyday literacy activities like reading, discussing, or thinking critically. Building on the situation model by van Dijk and Kintsch (1983), students may pose questions by relating narrative or expository text meanings to their background knowledge and figuring out what the text is about and how it is structured, and venture into deeper inferential or critical queries, thereby reflecting their mental processes during reading. By generating questions, students activate comprehension monitoring (Otero & Kintsch, 1992), which plays an important role in detecting inconsistencies in ideas, seeking alternative explanations, and integrating and synthesizing ideas from multiple contexts. Evaluating SGQs may illuminate the extent of intratextual connections students make within the text as well as extratextual links that extend beyond the text to students’ lived experience and world knowledge (Murphy et al., 2016). SGQ involves dynamic, fluid, and expansive thinking and a deep understanding of the text, encouraging numerous causal relationships that enact “rich, multilayered, and memorable” conceptualizations (Taboada & Guthrie, 2006, p. 7).

A typology of SGQs can be further enriched by inquiry-driven disciplinary literacy approaches that encourage students to think and talk like literary critics, scientists, historians, or mathematicians (Learned, 2018; Moje, 2008; Spires et al., 2016). Encouraging SGQ can help all students think about literacy in social and cultural contexts, examine multiple viewpoints, and ask complex, compelling questions about issues such as power imbalances in social class, race, ethnicity, gender, ability, and resource allocation (Stone, 2017; Street, 2005). These inquiry-driven approaches are inclusive and empowering, as they enable “access to literacy practices which might previously have excluded some children” (Stone, 2017, p. 33).

The Current Study

The purpose of the present study is to fill in the gaps about the extent to which young students’ ability to generate high-quality questions based on text reading is associated with their reading and writing abilities, attitudes toward literacy, and metacognition. The study further investigates whether the relationships vary by the type of text read, specifically between expository and narrative text types. The study was guided by the following questions:
1. To what extent are students’ reading and writing abilities associated with their ability to generate high-quality questions?
2. Do students’ attitudes toward reading and writing, interests in the text, and perceived understanding of the text further predict their SGQ ability beyond their reading and writing ability?
3. Do different text types moderate the predictive relationships?

Method

Participants

Eighty-nine participants (47% female) in Grades 5 and 6 (56% Grade 6) were recruited from three public schools within six classrooms in Toronto, Canada. The three schools were part of the Toronto District School Board, the largest school board in Canada, with a very diverse, multilingual population. According to the Toronto District School Board (2020), students with a primary language other than English comprised 38% of School A’s population, 85% of School B’s, and 14% of School C’s. Within this study’s sample, 22% of the participants were born outside of Canada (primarily in Asia, including Bangladesh, China, India, Japan, Pakistan, South Korea, and Sri Lanka). Approximately 11% self-reported that they had ever taken an English as a second language class. When asked to self-rate their proficiency on a 5-point Likert scale in each of the languages they can speak, 25% of students reported their proficiency in English as being lower than that in at least one language other than English.

Local contexts. In Canada, education policies and curricula are developed at the provincial level; therefore the teachers in this study follow the Ontario provincial curriculum. The Ontario language standards (Ontario Ministry of Education, 2006) state that Grade 6 students are expected to read for meaning using a variety of text and media types through summarizing, identifying main ideas and supporting details, making inferences and interpretations, connecting to own experiences, analyzing and evaluating texts, and identifying points of view. The language standards refer to questioning in a variety of categories at the Grade 5–6 level, including oral communication (active listening strategy, comprehension strategy, determining point of view) and reading (determining point of view, metacognitive strategy). More broadly, the document states, “In language, students are encouraged from a very early age to develop their ability to ask questions and to explore a variety of possible answers to those questions” (Ontario Ministry of Education, 2006, p. 29). While the curriculum makes mention of this questioning strategy in several places, there is a lack of research on how such strategies are implemented within language instruction in the classroom setting.

Measures

All measures were developed as part of a larger research project that assessed students’ oral language, literacy, and learning beliefs (Sinclair et al., 2021). The measures used in this study are a questionnaire that asks about students’ demographics and learning
orientation, one writing task, and two reading comprehension measures with an associated question-generating task each. Teachers administered all measures within classrooms.

**Learning orientation.** Before engaging in the reading and writing tasks, students were asked to reflect on their attitudes toward reading and writing by responding to the statements “I like to read” and “I like to write” using a 5-point Likert scale ranging from “Definitely not” to “Yes, definitely.”

**Writing Task.** Students were given a 3-minute video stimulus about children’s use of social media and were asked to provide a written response on “whether social media is good or bad for young people.” The topics and stimuli were chosen to enhance student engagement and interest. Students were instructed to provide supporting details and examples from their personal experience to substantiate their argument. They were given approximately 45 minutes to plan and write their responses. Responses were scored in four categories: idea development, organization and coherence, vocabulary and expression, and grammatical range and accuracy. Supplemental Appendix A contains an overview of writing scoring categories. Each scoring category was defined by how scores captured various aspects of “good” writing, including content, form, syntactic language use, and lexical language use. A 4-point scale scoring rubric for each category was developed through an extensive literature review and multiple rounds of pilot scoring, inter-rater reliability checks, and discussions. Five raters, all of whom actively participated in developing and refining the rubric, assessed students’ writing. Raters were paired and randomly assigned to independently score students’ writing using the rubric. The average score given by each rater pair formed the final score. Discrepancies larger than two points between paired raters were moderated by a third rater, who provided a final decision. Cohen’s Kappa coefficients ranged from 0.82 to 0.89. The overall writing score was calculated by summing the scores from the four categories.

**Reading comprehension.** Students were provided with two reading passages. The passages were adopted from previous administrations of the Grade 6 provincial reading assessment in Ontario and repurposed by the research team for this study. The first was a narrative biographical text called “Marilyn Bell and Her Historic Swim,” based on a young female athlete who was the first to swim across Lake Ontario. The text described numerous challenges faced by the 16-year-old and her eventual triumph as she swam more than 50 kilometers in just over 24 hours. The second passage was an expository text called “What Do Bats Eat?,” which described different types of bats based on their dietary preferences. The text explained the differences between insectivores, nectarivores, frugivores, carnivores, and sanguivores. Both texts were analyzed using TextEvaluator, an automated tool that assesses complexity characteristics of reading passages, developed by the Educational Testing Service (Sheehan et al., 2014). Although the expository text was longer, the two passages had similar syntactic complexity, academic vocabulary, and number of words per sentence.
After reading each passage, participants were asked to rate their perceived understanding of the text. The question “How well do you think you understand this text?” was answered using a 5-point Likert scale that ranged from “Not very well” to “Very well.” This was followed by another question asking students to rate their interest in the given passage. Students responded to the question “How interesting is this text?” on a 5-point Likert scale ranging from “Not at all interesting” to “Very interesting.”

Finally, students completed nine multiple-choice comprehension questions for each passage. The comprehension questions elicited three different skills: (a) explicit comprehension, based on questions whose answers were explicitly stated in the text (e.g., Where did Marilyn start her swim? What do nectarivores eat?); (b) inferential comprehension, based on questions whose answers were not explicitly stated but could be inferred from the text (e.g., What was the author’s purpose in writing this text? What would likely happen if there were no bats?); and (c) discourse-level comprehension that required a broader text-level understanding and more complex evaluative skills (e.g., How is the reading organized? What is the main idea of this whole writing?). Supplemental Appendix B contains the narrative and expository texts and reading comprehension questions provided.

**Student-generated questions.** Students were asked to generate three questions for each passage without prior instruction or training. For the narrative passage, students’ questions were directed to the author of the text with instructions as follows: “Imagine you are meeting the author of ‘Marilyn Bell and Her Historic Swim.’ Based on what you read, create three questions you would ask the author.” Students’ questions were directed to an expert for the expository text, with this scenario: “Imagine you are meeting a bat expert. Based on what you have learned from this reading, what else do you want to learn about this topic?”

**Analysis.** Our typology of SGQs was derived inductively through iterative content analysis of 482 questions generated across the narrative and expository passages. In the present study, the quality of SGQs was defined as students’ ability to critically appraise expository and narrative texts by asking questions that elicited responses of varying complexity and extended beyond the scope of the passage. The quality of students’ questions was assessed using a literacy-specific typology. While Bloom’s taxonomy (Bloom et al., 1956) and subsequent revisions (Krathwohl, 2002) have often been used to assess the quality of SGQs, we felt the nature of students’ questions is more suited to a fluid typology rather than a leveled hierarchy. Moreover, SGQs can be more widely understood when embedded in real-life activities such as reading, discussing, investigating, and critiquing.

The SGQ typology consists of three mental models, each with two subdomains on a continuum (from 1 to 6). The three mental models represent students’ incorporation of text processing, situational knowledge, and critical appraisal (Table 1). The Text Processing model involved literal, syntactic, and lexical processing, with descriptive questions as the first domain in the typology. Descriptive questions were asked to seek literal meanings and recall or review information, terminology, facts, events,
|        | 1                          | 2                          | 3                          | 4                          | 5                          | 6                          |
|--------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
| Mental model | Text-processing           | Situational               | Critical                  |                           |                           |                           |
| Domain  | Descriptive               | Connective                | Generative                | Constructive              | Integrative               |                           |
| **Narrative examples** | How old was she [Marilyn]? | Was the story based on a true story? | Why did you wright your story on Marilyn Bell? | Did she [Marilyn] only do this for popularity, or also to change the rights? | Was your purpose to make known that women can be just as athletic as men? |
| **Expository examples** | Have humans ever encountered a vampire bat? | Can bats eat sweets? | What is your favorite type of bat? | Does the food come up differently to different bats when they use Echolocation? | If vampire bat saliva can be used for medical reasons, can they also be animal medicines? | If some bats kill and eat insects is research conducted to find more ways bats can help kill pests? |
and contexts. Examples of descriptive questions include “How fast did Marilyn Bell swim this length?” and “How many kinds of bats are there?” Students’ questions were also considered descriptive if they were non-codable, such as ones written as statements (e.g., “I want to learn what bats eat”) or ones that were unintelligible (e.g., “do you kind Marilyn personaly?”). Cognitive questions within the Text Processing model are those students ask to clarify meaning, explore why and how things are organized into parts, and inquire about similarities and differences. Examples of cognitive questions include “Was the story based on a true story?” and “Why don’t all bats suck blood?”

The Situational model includes connective and generative questions. Connective questions demonstrated attempts to link textual information to prior experience, understand the author’s intent, or infer gaps in the text and story characters, including emotions or intentions (Janssen et al., 2012). Examples of connective questions include “Do you think Marilyn Bell was scared?” and “How many bats eat fruit?” Generative questions seek to hypothesize, extend the storyline, seek solutions to problems, provide alternative explanations, or ponder a theme or main idea. Examples of generative questions include “Why did Marilyn Bell chose to swim across Lake Ontario, not another destination?” and “What happens to vampire bats if they get too less or too much blood a day?”

The Critical model includes constructive and integrative questions. Constructive questions highlighted students’ attempts to critically reflect on biases within social or historical contexts (Norris & Phillips, 2003; Street, 2005), faulty reasoning, and invalid assumptions within specific disciplines or the credibility of authors (Moje, 2008; Shanahan & Shanahan, 2008; Spires et al., 2016). Examples of constructive questions include “Do you believe this achievement was a big impact on our societie?” and “Can you have a bat as a pet or would this be bad for them?” Integrative questions seek ways to integrate information from multiple realms or criteria (Murphy et al., 2016). Some examples include “Have you written any other biographies of people who were determined to change how people think about a topic?” and “Can bats eventually make mosquitos extinct?”

Five raters manually coded 482 student-generated questions. Two raters assessed each question independently using the six-domain typology with scores ranging from 1 to 6. In the first round of rating, raters agreed on 76% of questions (366) and Cohen’s Kappa was 0.66. When discrepancies were greater than two levels, team discussion ensued and raters reexamined their original ratings. The final score was calculated by averaging the adjusted scores from the two raters.

Sample sizes, possible ranges, means, and standard deviations of all variables included in this study are displayed in Table 2. The columns under “Overall” show the sum scores for narrative and expository passages. The next two sets of columns are passage-specific scores for each of the narrative and expository texts. The writing ability and attitude toward reading and writing were not text-specific, so these variables have one overall value for each. Supplemental Appendix C provides the correlation matrix for the variables used in the subsequent analyses.

A hierarchical regression model was used with the sum score of the quality of SGQs as the outcome variable. A sample size of 89 was deemed adequate for a regression
Table 2. Sample Sizes, Possible Range, Means, Standard Deviations of All Variables.

| Variable set     | Variable               | Overall |           |          | Narrative text |           |          | Expository text |           |          |
|------------------|------------------------|---------|-----------|-----------|----------------|-----------|-----------|----------------|-----------|-----------|
|                  |                        | n    | Possible range | M   | SD  | n    | Possible range | M   | SD  | n    | Possible range | M   | SD  |
| Outcome          | Quality of SGQ         | 70   | 0–36      | 19.0 | 3.7 | 79   | 0–18      | 10.2 | 2.4 | 80   | 0–18      | 8.7  | 2.7 |
| Ability          | Reading comprehension  | 76   | 0–18      | 13.5 | 2.2 | 85   | 0–9       | 6.5  | 1.3 | 79   | 0–9       | 7.0  | 1.4 |
|                  | Writing                | 81   | 4–16      | 9.6  | 2.7 | -    | -         | -   | -   | -    | -         | -    | -   |
| Attitude         | Attitude toward reading| 81   | 1–5       | 4.2  | 1.1 | -    | -         | -   | -   | -    | -         | -    | -   |
|                  | Attitude toward writing| 81   | 1–5       | 3.6  | 1.2 | -    | -         | -   | -   | -    | -         | -    | -   |
| Meta-cognition   | Perceived understanding| 78   | 2–10      | 8.9  | 0.8 | 86   | 1–5       | 4.5  | 0.5 | 80   | 1–5       | 4.3  | 0.5 |
|                  | Interest in passage    | 78   | 2–10      | 7.8  | 1.3 | 86   | 1–5       | 4.0  | 0.7 | 80   | 1–5       | 3.9  | 0.7 |

Note. \(N=89\). The possible range is the possible range each student could score within each variable.
model with six independent variables (Tabachnick & Fidell, 2001). A series of regression diagnostics were performed for assumption checks with three sets of variables—(a) two texts combined, (b) narrative text, and (c) expository text. Multicollinearity was examined; none of the independent variables were too highly correlated, and the collinearity statistics were acceptable (VIF = 1.1–1.6). Scatter plots between variables, P–P plots, and Q–Q plots were also examined for the assumptions of linearity, normality, and homoscedasticity, which indicated no concerns of assumption violation. The results from the Shapiro–Wilk W test of normality \( p > .05 \) and the White test for homoscedasticity \( p > .05 \) further confirmed that the data met these assumptions.

We used the multiple imputation technique to prevent any loss of statistical power caused by missing data (Rubin, 1987). In the observed data, 3%–12% of the values were missing in the variables used in the analysis. Most students with missing values randomly skipped some questions, while others could not complete some portions of the measures due to their absence from school during the administration of those measures. A series of cross-tabulations, logistic regressions, and \( t \)-tests were conducted to investigate any missing data patterns. The results indicated that it is reasonable to assume the missing data mechanism is missing at random rather than missing not at random (Schafer & Graham, 2002). All variables that are of interest in this study were used in the imputation phase, including the outcome variable (the quality of SGQ), as supported by Graham (2009). Because there was no demographic variable that predicted the missingness of the data, no auxiliary variables were added. Ten imputed data sets were created using chained equations (MICE) through the mi suite of commands in Stata 16 (StataCorp., 2019). The distributions of completed data were compared graphically to those of the observed data (Eddings & Marchenko, 2012). The performance of the imputation model was considered acceptable, as no major differences in the data distribution were found in these diagnostics.

A three-step hierarchical regression was conducted with the quality of SGQ variable as the dependent variable. The reading and writing ability variables were entered in the first step to control for students’ literacy ability. Two attitudinal variables were entered at step 2 and two metacognitive (perceived understanding and interest) variables at step 3. We compared the overall effects of both the narrative and expository results combined, as well as each genre separately. Analyses were conducted using Stata 16 (StataCorp., 2019). An alpha level of 0.05 was used for all statistical tests.

**Results**

Results of the three-step hierarchical regression analyses are summarized in Table 3. In step 1, reading ability significantly predicted the quality of SGQ, \( t(86) = 2.56, p = .01 \), with the two variables accounting for 13% of the variance in overall SGQ. When text types were considered, reading ability was not a statistically significant predictor of SGQ quality for the narrative text. On the other hand, approximately 11% of the SGQ variance for the expository text was accounted for by reading ability, \( t(86) = 2.46, p = .02 \). Interestingly, qualitative examination of SGQs revealed many lower performing students were still able to produce higher complexity critical thinking
Table 3. Hierarchical Regression Analysis for Prediction of the Quality of SGQ.

| Variable                      | Two passages combined | Narrative text | Expository text |
|-------------------------------|-----------------------|----------------|-----------------|
|                               | R² | ΔR² | β   | p | R² | ΔR² | β   | p | R² | ΔR² | β   | p |
| Step 1: Ability               |    |     |     |    |    |     |     |    |    |    |     |     |    |
| Reading score                 | .13 | -  | .11 | .43 | .35 | .02*|    |    |    |    |     |     |    |
| Writing score                 | .05 | .80 | .13 | .33 | −.07 | .62|    |    |    |    |     |     |    |
| Step 2: adding Attitude       |    | .11 | .19 | .14 | .14 | .03|    |    |    |    |     |     |    |
| Reading score                 | .47 | .03*|     | .05 | .69 | 0.33| .02*|    |    |    |     |     |    |
| Writing score                 | 0.00 | .98 | 0.18 | .39 | −0.09 | .51|    |    |    |    |     |     |    |
| I like to read                | 0.38 | .33 | 0.00 | .99 | 0.16 | .21|    |    |    |    |     |     |    |
| I like to write               | 0.86 | .01**| 0.38 | .00**| 0.05 | .68|    |    |    |    |     |     |    |
| Step 3: adding Meta-cognition |    | .07 | .25 | .06 | .16 | .02|    |    |    |    |     |     |    |
| Reading score                 | 0.42 | .05*| 0.05 | .70 | 0.31 | .03*|    |    |    |    |     |     |    |
| Writing score                 | 0.06 | .70 | 0.08 | .49 | −0.08 | .56|    |    |    |    |     |     |    |
| I like to read                | 0.30 | .44 | −0.06 | .61 | 0.16 | .22|    |    |    |    |     |     |    |
| I like to write               | 0.73 | .02*| 0.42 | .00**| 0.04 | .74|    |    |    |    |     |     |    |
| Perceived understanding       | 1.09 | .03*| 0.22 | .07 | 0.12 | .32|    |    |    |    |     |     |    |
| Interest in passage           | −0.49 | .13 | −0.21 | .11 | 0.02 | .88|    |    |    |    |     |     |    |

Note. *p < .05; **p < .01; ***p < .001.
questions. For example, students who scored less than 4 out of 9 (the lowest 10%) on the multiple-choice comprehension questions from the narrative text sometimes generated insightful questions directed to the passage author, such as “Why did she want to swim across Lake Ontario?,” “Is this story something that inspired you to write about?,” and “What are your sources?” Similarly, some students who performed the lowest in understanding expository texts also generated thoughtful questions, such as “Why don’t all bats suck blood?,” “How many types of bats are there?,” and “What is the most common type [of bat]?”

Introducing two attitudinal variables in step 2 accounted for an additional 11% of the total variance in the overall SGQ quality rating. In addition to reading ability, positive attitude toward writing was a statistically significant predictor of SGQ quality, $t(84) = 2.77, p = .01$, while the reading attitude variable was not. A similar result was observed with the narrative text type in that the attitudinal variables accounted for an additional 14% of the total SGQ variance associated with the narrative text type. It was mostly the writing attitude variable that contributed to the additional $R^2$, $t(84) = 3.34, p = .00$, and reading ability no longer significantly predicted the quality of SGQ for the narrative text. The two attitudinal variables increased the $R^2$ by only 3% for the expository text, and none of these variables were statistically significant.

Finally, the addition of two metacognitive variables (perceived textual understanding, interest in the text topic) to the regression model explained an additional 7% of the total SGQ variance. Perceived understanding of the text was a statistically significant predictor, $t(82) = 2.20, p = .03$, in addition to reading ability and attitude toward writing. Together, three independent variables, reading ability, writing attitude, and perceived understanding, accounted for approximately 31% of the variance in SGQ. When the addition of these metacognitive variables to the regression model was examined by text type, neither was a statistically significant predictor of SGQ quality.

**Discussion**

In the present study, we examined the extent to which students’ abilities, attitudes, perceived understanding, and interests in reading and writing predict the quality of SGQs between text genres. Overall, the quality of SGQ was significantly predicted by reading comprehension, positive attitude toward writing, and students’ perceived understanding of the text. However, writing ability, attitude toward reading, and interest in the topic do not seem to be significant predictors of SGQ quality.

The strong association between SGQ quality and reading comprehension supports previous research findings that SGQ activates heightened levels of attention and text engagement (Kraal et al., 2018; Taboada & Guthrie, 2006). Some studies posit that more engaged learners are intrinsically motivated, aspire to more complex skills and abilities (Guthrie & Cox, 2001), and incorporate more complex metacognitive strategies, which result in deeper, more meaningful understanding (Singer & Donlan, 1982). Conversely, Davey and McBride (1986) reported that reading ability was not significantly related to the quality of question generation. The present study corroborates this unusual result since qualitative data inspection revealed that some lower performing students generated high-quality questions. This suggests that SGQ ability may
transcend literacy ability. This can have profound implications for learning and assessment of all learners. For example, this association may suggest a troubling disparity between what traditional reading comprehension assessments may capture and what students can actually do above and beyond understanding written texts. Therefore, using SGQ with a variety of genres in assessments may produce a more authentic and holistic picture of what diverse students know and understand while enhancing valuable student agency within the process.

In addition, this study showed a statistically significant, positive relationship between SGQs and students’ attitude toward writing as well as self-perceived understanding of the text. The result that SGQ quality is associated with positive attitude toward writing suggests the expressive nature of articulating queries that SGQ elicits. This may further suggest that metacognitive and attitudinal variables in learning play a larger role than was previously thought and warrant further investigation (Guthrie & Cox, 2001).

Given the similarities between writing processes and question generation, this study aimed to examine whether writing ability predicts SGQ quality. Although our results were not statistically significant, null findings may not necessarily allude to a null relationship but may in fact call for a deeper examination of the relationship between specific aspects of writing processes and question generation. For example, the idea-development aspect of writing may reveal stronger associations with SGQ quality because a student may be very skilled at generating ideas but may struggle with organizing and expressing those ideas with appropriate vocabulary and writing conventions.

Another predictor of SGQ quality was students’ self-rated level of understanding, elicited by the question “How well do you think you understand this text?” This response captured the metacognitive dimension of students’ learning. A heightened self-awareness of comprehension ability seems to build confidence and perhaps a firmer understanding of the text, resulting in higher quality questions (Davey & McBride, 1986).

Once the data were separated by narrative and expository text, there appeared to be a strong genre effect impacting the quality of SGQs, where different variables had varying effects across genres. High-quality questions in the narrative text were associated with a positive attitude toward writing, whereas in the expository text, the strongest predictor was reading comprehension ability, or how well students understood the text. It is noteworthy that no predictor was significant across both genres. Therefore, it is possible that different genres provoke different types of responses and information processing, as is consistent with current research (Bugg & McDaniel, 2012; Kraal et al., 2018; Taboada et al., 2012). This suggests that SGQ may serve different purposes, calling for different pedagogical approaches within different learning and assessment contexts.

Another interesting area in need of further examination is the impact of training students to generate good questions. Numerous studies confirm that students who are taught how to ask critical questions tend to ask questions that involve more complex processes such as extending and expanding conceptual relationships in texts to broader contexts (Davey & McBride, 1986; Taboada et al., 2012). Students’ literal
and inferential reading comprehension improve when they are taught to ask specific questions, generate question stems, and identify genre-specific text structures (Baleghizadeh & Babapour, 2011; Taboada & Guthrie, 2006). Although the current study does not include a student instructional component, further research is needed about how this pedagogical strategy may benefit student agency, learning, and authentic assessment.

Existing literature has also revealed significant variations in the kinds of learning opportunities students are exposed to within classrooms. For example, Mitani (2021) explained that aspects such as low socioeconomic status or ethnic and racial minorities affect how teachers perceive student ability levels, whereby these students often end up in “basic classes” that focus on “mastery of low-level skills” (p. 1), which may limit opportunities to engage in more complex learning and thinking activities. Keehne et al. (2018) illustrated the dynamic impacts of culturally responsive instruction, where student-led cultural inquiry perpetuates traditional language revitalization, cultural identity, and community advocacy, while simultaneously building higher level thinking skills and literacy confidence among diverse student ability levels. While socioeconomic status and cultural variances are beyond the scope of this study, SGQ may help level the playing field, where students of all abilities and backgrounds can have the opportunity to inquire, discuss, and grapple with deep, meaningful ideas. Moreover, a more inclusive and flexible approach to conceptualizing student thinking, learning, and questioning may contribute to expanding the playing field, as the current study attempts to do. Additional research is needed that explores the impacts of sociocultural aspects in learning, thinking, and questioning.

**Implications**

The study results provide important insights into some of the factors associated with question-generation effects and the impact on learning and assessment. Overall, significant predictors of SGQ quality are reading comprehension, positive attitudes toward writing, and students’ perceived level of text understanding. SGQ’s overall positive association with reading comprehension suggests expanded use of SGQ in classrooms could be beneficial. Teachers can include SGQ as a pedagogical strategy to improve students’ comprehension of expository text. This can be especially useful for lower level comprehenders and younger students since the use of expository texts in school seems to increase dramatically after Grade 4, when students move from using mostly narrative texts to working with more complex and less familiar expository texts (Kraal et al., 2018). For example, Roehling et al. (2017) cited the determination of text structure as a challenging aspect of expository text comprehension and suggested that students ask themselves “guiding questions” to help focus on structure-related elements (p. 74). Teachers can also incorporate SGQ into language arts activities to encourage close, critical reading on a deeper, more meaningful level (Fisher & Frey, 2015). Indeed, SGQ in any discipline could improve students’ understanding of the content. It can uncover fascinating quandaries or elusive mysteries that students can dig into, discuss, and research. Moreover, SGQ could be used in a group dynamic where students discuss questions together and co-construct meaning through
negotiated understandings (Chin & Brown, 2002). Regardless, SGQ appears to spark autonomy and intrinsic motivation and bring learner-centeredness alive. As Chin and Brown (2002) confirmed, a hallmark of “self-directed, reflective learners is their ability to ask themselves questions that help direct their learning” (p. 522).

The overall positive association of SGQ quality with students’ attitudes toward writing and perceived understanding may potentially enhance agency and confidence for all learners, and facilitate a more authentic, student-centered approach to learning. Moreover, the possibility that SGQ may sometimes transcend literacy ability could provide a welcome boost of confidence for students who are experiencing challenges, thereby increasing academic self-efficacy. As Stone (2017) suggested, all students, regardless of age and ability, might possess a “natural ability to ask critical questions” (p. 55). Therefore, SGQ could be used as a pedagogical intervention strategy to create more active reading and foster critical thinking skills for students of all ability levels, especially those with lower literacy levels (Davey & McBride, 1986), English language learners (Taboada et al., 2012), or even emergent readers (Stone, 2017). The positive attitude toward writing as a predictive factor for SGQ quality also suggests that SGQ activities could potentially be used to help students generate ideas for writing composition or research initiatives. However, given the statistically insignificant association between writing ability and SGQ quality, additional support for organization and expression may need to be provided along with SGQ activities to develop student writing skills. Further research in these areas would be beneficial.

SGQ can also be used creatively within a variety of genres to help address the disparity between what traditional assessments may capture and what students actually know and understand. For example, students can generate interest-based queries on course material, seek answers, and discuss their responses as they demonstrate what they have learned. This moves students beyond memorization to more complex cognition and places them as “active agents” in meaningful learning (Jones, 2019). As Guthrie and Cox (2001) surmised, evaluating students’ understanding through their ability to solve problems and seek meaningful responses to their questions “is a more aligned form of assessment than answering disconnected true and false questions” (p. 294). SGQs act as a formative assessment that can reveal misunderstandings or knowledge gaps where teachers can provide additional opportunities for students to construct meaning (Chin & Brown, 2002). Although not within the scope of this study, it would be interesting to explore the use of SGQ as a diagnostic tool to gauge students’ understanding of concepts as they unfold in the classroom.

The contrasting genre-specific effects that emerged also have implications for learning and assessment. In narrative texts, the association between SGQ quality and a positive attitude toward writing means that metacognitive aspects such as learning attitude and emotion may have implications that extend beyond SGQ. SGQ can be used to strengthen low performers’ academic self-efficacy, enhance agency and control, and contribute to overall academic success. In addition, SGQ activities can be used with narrative texts to help students generate ideas for writing composition in language arts. For expository texts, the sole significant finding of reading comprehension on SGQ quality suggests that SGQ in expository contexts may have more benefit for higher performing students. It appears that the complexity of students’ questions
increases as their understanding of a topic increases. As Miyake and Norman (1979) pointed out, more knowledgeable people tend to ask more critical, evaluative-type questions. While the present study’s genre-specific findings support previous evidence that training in SGQ improves reading comprehension scores (Taboada & Guthrie, 2006), some lower performing students, without any training in SGQ, were still able to generate high-quality questions in both narrative and expository texts. This highlights the troubling inconsistency of what traditional reading comprehension assessments may measure and what students actually understand when working with diverse text types.

Conclusion

While the implications of this study are many, we acknowledge some methodological limitations. One limitation may be the number of questions included in each passage that were used to measure students’ reading comprehension ability. More questions per construct could help to uncover latent variables that may further impact SGQ quality. Similarly, the use of only two texts, one narrative and one expository, is a limiting factor. A wider range of text genres and more text samples would create more opportunities for detailed comparisons. Another limitation is the relatively small sample size of this study. A larger sample would provide more statistical power and allow for more precise estimates. While the study took place in a diverse, multicultural city, additional work that more specifically tracks and targets known correlates of academic achievement, such as socioeconomic status, would be beneficial (Bradley et al., 2001; Mitani, 2021). One final limitation is that no genre-specific writing ability scores were available, due to the methodological constraints of the study and the fact that the writing task did not match up with the genres used in the reading comprehension task, thereby limiting comparability.

Regardless of the approach or context, it seems clear that SGQ has the potential to significantly enhance literacy learning, pedagogy, and assessment. What is particularly noteworthy is that these positive impacts result from a fairly brief and generalizable activity that provides transferrable skills applicable to any lesson or assessment and can be beneficial to a wide range of learners. While additional benefits may be gained from adding a training component to SGQ, the results we have gathered in this study spawn from the straightforward act of generating one’s own questions. Moreover, a teacher’s or researcher’s carefully crafted questions to the student can guide the student’s attentional focus and perhaps further influence the positive impact of the SGQ (Taboada & Guthrie, 2006), yet the agency remains in the hands of the student—where it belongs! In order for research to have a true impact, it is important that researchers strive to disseminate results in ways that reach right into classrooms where knowledge is co-created and transformed.

More research is needed to elucidate the use of SGQs in literacy learning and assessment. For example, how might our SGQ findings manifest in emergent readers? What are the implications of SGQ with English language learners or those with specific learning disabilities? Do other factors, such as working memory, vocabulary, gender, socioeconomic status, or cultural variations, influence SGQ quality? The use
of critical literacy and media texts as constructs would be useful to further tease out the predictive aspects of SGQ quality across a range of mediating and moderating factors. Introducing collaborative group discussions and co-construction of knowledge may help shed further light onto the benefits of SGQ. Additional research on how SGQ impacts the fostering of constructive, integrative thinking for diverse students in diverse contexts is needed. In addition, studies with larger sample sizes are needed to expand statistical power. Finally, incorporating student choice in the methodology regarding which texts they interact with when generating their questions could help to further investigate the motivational dimensions of SGQ and learning (Taboada et al., 2012).

While Betts (1910) stated many years ago that the “skill in the art of questioning lies at the basis of all good teaching” (as cited in Chin & Brown, 2002, p. 547), one could add that the skill in the art of questioning lies at the basis of all good teaching and assessment. Yet, given the skills that twenty-first-century learners require in today’s realities, and the active student-centeredness that lives at the heart of learning, perhaps we must heed Chin and Brown’s (2002) more relevant proposition that “to know how to question is to know how to learn well” (p. 547). Indeed, as Eisner (2001) posited, “we should be less concerned with whether [students] can answer our questions than whether they can ask their own” (p. 370).

**Declaration of Conflicting Interests**

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The authors received no financial support for the research, authorship and/or publication of this article.

**ORCID iD**

Lois Maplethorpe https://orcid.org/0000-0001-6290-7917

**Supplemental Material**

The appendix referenced in this article and abstracts in languages other than English are available at: http://journals.sagepub.com/doi/suppl/10.1177/1086296X221076436.

**References**

Au, K. H. (1980). Participation structures in a reading lesson with Hawaiian children: Analysis of a culturally appropriate instructional event. *Anthropology and Education, 11*(2), 91–115. https://doi.org/10.1525/aeq.1980.11.2.05x1874b

Baleghizadeh, S., & Babapour, M. (2011). The effect of summary writing on reading comprehension and recall of EFL students. *New England Reading Association Journal, 47*(1), 44–48.
Bates, S. P., Galloway, R. K., Riise, J., & Homer, D. (2014). Assessing the quality of a student generated question repository. Physical Review Special Topics: Physics Education Research, 10(2), 1–11. https://doi.org/10.1103/PhysRevSTPER.10.020105

Becker, R. R. (2000). The critical role of students’ questions in literacy development. The Educational Forum, 64(3), 261–271. https://doi.org/10.1080/00131720008984763

Betts, G. H. (1910). The recitation. Houghton Mifflin.

Bloom, B. S., Englehart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). Taxonomy of educational objectives: Handbook I. Cognitive domain. McKay.

Botsas, G., & Padeliadu, S. (2003). Goal orientation and reading comprehension strategy use among students with and without reading difficulties. International Journal of Educational Research, 39(4–5), 477–495. https://doi.org/10.1016/j.ijer.2004.06.010

Bradley, R. H., Corwyn, R. F., McAdoo, H. P., & Garcia Coll, C. (2001). The home environments of children in the United States part I: Variations by age, ethnicity, and poverty status. Child Development, 72(6), 1844–1867. https://doi.org/10.1111/1467-8624.t01-1-00382

Bugg, J. M., & McDaniel, M. A. (2012). Selective benefits of question self-generation and answering for remembering expository text. Journal of Educational Psychology, 104(4), 922–931. https://doi.org/10.1037/a0028661

Cain, K., Oakhill, J., & Bryant, P. (2000). Investigating the causes of reading comprehension failure: The comprehension-age match design. Reading and Writing: An Interdisciplinary Journal, 12(1–2), 31–40.

Cazden, C. B. (1988). Classroom discourse: The language of teaching and learning. Pearson Education Canada.

Chin, C., & Brown, D. E. (2002). Student-generated questions: A meaningful aspect of learning in science. International Journal of Science Education, 24(5), 521–549. https://doi.org/10.1080/09500690110095249

Cohen, R. (1983). Self-generated questions as an aid to reading comprehension. The Reading Teacher, 36(8), 770–775. https://www.jstor.org/stable/20198324

Davey, B., & McBride, S. (1986). Effects of questioning-generation training on reading comprehension. Journal of Educational Psychology, 78(4), 256–262. https://doi.org/10.1037/0022-0663.78.4.256

Demirdogen, B., & Cakmakci, G. (2014). Investigating students’ interest in chemistry through self-generated questions. Chemistry Education Research and Practice, 15(2), 192–206. https://doi.org/10.1037/C4RP00037D

Duncan, T., & Buskirk-Cohen, A. (2011). Exploring learner-centered assessment: A cross-disciplinary approach. International Journal of Teaching and Learning in Higher Education, 23(2), 246–259.

Eddings, W., & Marchenko, Y. (2012). Diagnostics for multiple imputation in Stata. The Stata Journal, 12(3), 353–367. https://doi.org/10.1177/1536867X1201200301

Eisner, E. W. (2001). What does it mean to say a school is doing well? Phi Delta Kappan, 82(5), 367–372. https://doi.org/10.1177/001317210108200506

Feldt, R. C., Feldt, R. A., & Kilburg, K. (2002). Acquisition, maintenance, and transfer of a questioning strategy in second- and third-grade students to learn from science textbooks. Reading Psychology, 23(3), 181–198. https://doi.org/10.1080/02702710290061319

Fisher, D., & Frey, N. (2015). Teacher modeling using complex informational texts. The Reading Teacher, 69(1), 63–69. https://doi.org/10.1002/trtr.1372

Flower, L., & Hayes, J. (1981). The pregnant pause: An inquiry into the nature of planning. Research in the Teaching of English, 15(3), 229–243. http://www.jstor.org/stable/40170791
Gonzales, N. A. (1996). Problem formulation: Insights from student generated questions. *School Science and Mathematics, 96*(3), 152–157. https://doi.org/10.1111/j.1949-8594.1996.tb15830.x

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60*(1), 549–576. https://doi.org/10.1146/annurev.psych.58.110405.085530

Guthrie, J. T., & Cox, K. E. (2001). Classroom conditions for motivation and engagement in reading. *Educational Psychology Review, 13*(3), 283–302. https://doi.org/10.1023/A:1016627907001

Guthrie, J. T., & Wigfield, A. (2000). Engagement and motivation in reading. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (pp. 403–422). Routledge.

Holliday, W. G., Yore, L. D., & Alvermann, D. E. (1994). The reading-science learning-writing connection: Breakthroughs, barriers, and promises. *Journal of Research in Science Teaching, 31*(9), 877–896. https://doi.org/10.1002/tea.3660310905

Jang, E. E. (2014). *Focus on assessment*. Oxford University Press.

Janssen, T., Braaksma, M., Rijlaarsdam, G., & van den Bergh, H. (2012). Flexibility in reading literature: Differences between good and poor adolescent readers. *Scientific Study of Literature, 2*(1), 83–107. https://doi.org/10.1075/ssol.2.1.05jan

Jones, J. (2019). Scaffolding self-regulated learning through student-generated quizzes. *Active Learning in Higher Education, 20*(2), 115–126. https://doi.org/10.1177/1469787417735610

Keehne, C. N. K., Sarsona, M. W., Kawakami, A. J., & Au, K. (2018). Culturally responsive instruction and literacy learning. *Journal of Literacy Research, 50*(2), 141–166. https://doi.org/10.1177/1086296X18767226

Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. Cambridge University Press.

Kraal, A., Koormeef, A. W., Saab, N., & van den Broek, P. W. (2018). Processing of expository and narrative texts by low- and high-comprehending children. *Reading and Writing, 31*(9), 2017–2040. https://doi.org/10.1007/s11145-017-9789-2

Krathwohl, D. R. (2002). A revision of Bloom’s taxonomy: An overview. *Theory Into Practice, 41*(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2

Learned, J. (2018). Doing history: A study of disciplinary literacy and readers labeled as struggling. *Journal of Literacy Research, 50*(2), 190–216. https://doi.org/10.1177/1086296X17746446

Levin, T., & Wagner, T. (2005). Enhancing thinking dispositions through informal writing: Experiences in science classes. In G. Rijlaarsdam, H. van den Bergh, & M. Couzijn (Eds.), *Studies in writing: Vol. 14. Effective learning and teaching of writing* (2nd ed., pp. 481–4978). Kluwer Academic Publishers.

Mitani, H. (2021). Test score gaps in higher order thinking skills: exploring instructional practices to improve the skills and narrow the gaps. *AERA Open, 7*(1), 1–23. https://doi.org/10.1177/23328584211016470

Miyake, N., & Norman, D. A. (1979). To ask a question, one must know enough to know what is not known. *Journal of Verbal Learning and Verbal Behavior, 18*(3), 357–364. https://doi.org/10.1016/S0022-5371(79)90200-7

Moje, E. B. (2008). Foregrounding the disciplines in secondary literacy teaching and learning: A call for change. *Journal of Adolescent & Adult Literacy, 52*(2), 96–107. https://doi.org/10.1598/IAAL.52.2.1

Murphy, P. K., Andiliou, A., Firetto, C. M., Bowersox, C. M., Baker, M., & Ramsay, C. M. (2016). Intratextual persuasive messages as catalysts for higher order thinking: An
exploratory investigation. *Journal of Literacy Research*, 48(2), 134–163. https://doi.org/10.1177/1086296X16660652

Norris, S. P., & Phillips, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education*, 87(2), 224–240. https://doi.org/10.1002/sce.10066

Ontario Ministry of Education. (2006). *The Ontario curriculum, Grades 1-8: Language*. http://www.edu.gov.on.ca/eng/curriculum/elementary/language18currb.pdf

Organisation for Economic Co-operation and Development. (2019). PISA 2018 Assessment and analytical framework. In *PISA 2018 Assessment and analytical framework* (pp. 21–71). OECD Publishing. https://doi.org/10.1787/5c07e4f1-en.

Osborne, J. (2006). Message from the president. *E-NARST News*, 49(2), 1–20. https://narst.org/sites/default/files/2019-08/e-narstnews_2006_jul.pdf

Otero, J., & Kintsch, W. (1992). Failures to detect contradictions in a text: what readers believe versus what they read. *Psychological Science*, 3(4), 229–236. https://doi.org/10.1111%2Fj.1467-9280.1992.tb00034.x

Oxford, R. (1990). *Language learning strategies*. Heinle and Heinle.

RAND Reading Study Group. (2002). *Toward an R&D program in reading comprehension*. RAND. http://www.rand.org/pub

Resnick, L. B. (1987). *Education and learning to think*. National Academies Press.

Roehling, J. V., Hebert, M., Nelson, R. J., & Bohaty, J. J. (2017). Text structure strategies for improving expository reading comprehension. *The Reading Teacher*, 71(1), 71–82. https://doi.org/10.1002/trtr.1590

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. Wiley.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7(2), 147–177. https://doi.org/10.1037%2F1082-989X.7.2.147

Shanahan, T., & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy. *Harvard Educational Review*, 78(1), 40–59. https://doi.org/10.17763/haer.78.1.v62444321p602101

Sheehan, K. M., Kostin, I., Napolitano, D., & Flor, M. (2014). The TextEvaluator tool: Helping teachers and test developers select texts for use in instruction and assessment. *Elementary School Journal*, 115(2), 184–209. https://doi.org/10.1086/678294

Sinclair, J., Jang, E. E., & Rudzicz, F. (2021). Using machine learning to predict children’s reading comprehension from linguistic features extracted from speech and writing. *Journal of Educational Psychology*, 113(6), 1088–1106. https://doi.org/10.1037/edu0000658

Singer, H., & Donlan, D. (1982). Active comprehension: Problem-solving schema with question generation for comprehension of complex short stories. *Reading Research Quarterly*, 17(2), 166–185. https://doi.org/10.2307/747482

Spires, H. A., Kerkhoff, S. N., & Graham, A. C. (2016). Disciplinary literacy and inquiry: Teaching for deeper content learning. *Journal of Adolescent & Adult Literacy*, 60(2), 151–161. https://doi.org/10.1002/jaal.577

StataCorp (2019). *Stata (Release 16)* [Statistical software].

Stone, K. (2017). *Reconsidering primary literacy: enabling children to become critically literate*. Taylor & Francis.

Street, B. (2005). At last: Recent applications of new literacy studies in educational contexts. *Research in the Teaching of English*, 39(4), 417–423. http://www.jstor.org/stable/40171646

Tabachnick, B. G., & Fidell, L. (2001). *Using multivariate statistics*. Allyn and Bacon.

Taboada, A., Bianco, S., & Bowerman, V. (2012). Text-based questioning: A comprehension strategy to build English language learners’ content knowledge. *Literacy Research and Instruction*, 51(2), 87–109. https://doi.org/10.1080/19388071.2010.522884
Taboada, A., & Guthrie, J. T. (2006). Contributions of student questioning and prior knowledge to construction of knowledge from reading information text. *Journal of Literacy Research, 38*(1), 1–35. https://doi.org/10.1207/s15548430jlr3801_1

Torgesen, J. K., & Hudson, R. F. (2006). Reading fluency: critical issues for struggling readers. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 130–158). International Reading Association.

Toronto District School Board. (2020). *Facts and figures*. https://www.tdsb.on.ca/Find-your/Schools/Facts-and-Figures/schno/4208

Unrau, N., Ragusa, G., & Bowers, E. (2014). Teachers focus on motivation for reading: “It’s all about knowing the relationship.”. *Reading Psychology, 36*(2), 105–144. https://doi.org/10.1080/02702711.2013.836582

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. Academic Press.

Verlaan, W., Ortlieb, E., & Oakes-Verlaan, S. (2014). Developing argument writing through evidence-based responses to student-generated questions. *Voices From the Middle, 21*(3), 34–40.

**Manuscript submitted:** 9 December 2020

**Manuscript accepted:** 1 December 2021

**Article published online:** 10 February 2022