Making Sense of Science Texts: 
A Mixed-Methods Examination of Predictors and Processes of Multiple-Text Comprehension

Dennis S. Davis  
North Carolina State University, Raleigh, USA

Becky Huang  
Tanisha Yi  
The University of Texas at San Antonio, USA

ABSTRACT
Previous research has identified various factors that contribute to readers’ comprehension of expository texts, including strategy expertise, language proficiency, prior knowledge, and more recently, readers’ beliefs about knowledge. This study addresses the need to understand the relative contributions of these predictors to readers’ comprehension of multiple texts and the processes used by readers to make sense of texts. Eighty-three students (grades 5-7) participated in this mixed-methods study. The sample consisted of monolingual students and emergent and proficient bilingual students who completed measures of expository comprehension, strategic knowledge and awareness, English-language proficiency, prior content knowledge, and epistemic beliefs. Ten bilingual students from this sample also completed a think-aloud protocol to allow for close examination of their meaning-making processes. In a multiple regression analysis, English-language proficiency was the strongest predictor of comprehension, followed by content knowledge. Strategy knowledge and awareness and epistemic beliefs were not related to multiple-text comprehension in the model. The relationship between English-language proficiency and comprehension was stronger for bilingual students than for monolingual students. Students in the think-aloud sample demonstrated emergent knowledge of processes of disciplinary reading of multiple texts, including metacognitive monitoring, theorizing authorial identity, and intertextual integration, while also displaying a tendency to defer to institutionalized authority when evaluating credibility of the texts. The findings provide directions for future research on the way young adolescents comprehend and learn from expository texts in the discipline of science.

Expertise in reading comprehension is crucial for accessing discipline knowledge in school, particularly in science classrooms where content is often communicated through complex texts. Students need to be able to expertly handle the demands placed on them by science textbooks and other scientific sources to be fully apprenticed into the culture and practices of scientific knowledge creation (Osborne, 2014; Pearson, Moje, & Greenleaf, 2010). However, comprehension of science text is no simple feat. Science texts often contain advanced academic vocabulary, require readers to make complex inferences, and place demands on readers’ prior knowledge (Best, Rowe, Ozuru, & McNamara, 2005). Science texts also rely on discourse patterns that distinguish them from texts in other disciplines (Fang, 2006).
The challenges associated with science reading can be particularly daunting for readers who are in the process of developing proficiency in English (Fradd, Lee, Sutman, & Saxton 2001). Given that over 4 million (9.2%) K–12 students are currently identified as English learners (ELs) in public schools in the United States (Kena et al., 2015), there is a need to better understand the predictors and processes of reading comprehension of science texts among linguistically diverse students.

Previous Research

In the following sections, we briefly review the literature on the relationships between reading comprehension and the four predictors of interest in this study: strategic expertise, language proficiency, content knowledge, and epistemological beliefs. We discuss research conducted with both monolingual and bilingual students. Although there are other factors that could also be considered in our model, these four predictors were selected as the foci of the present study because they are all theorized to make important independent contributions to reading comprehension and have not been frequently studied together. Strategy expertise, language proficiency, and prior content knowledge have been heavily emphasized in research-based comprehension instruction (Duke, Pearson, Strachan, & Billman, 2011), including instruction for ELs (Baker et al., 2014). Recent conceptualizations of expository reading practices have suggested that epistemological beliefs (about knowledge and its construction) might also influence comprehension, particularly when readers are constructing integrated understandings across multiple disciplinary texts (Barzilai & Eshet-Alkalai, 2015; Goldman & Scardamalia, 2013).

Cognitive and Metacognitive Strategies for Meaning Making

Expertise in the use of meaning-making strategies has been found to correlate with reading comprehension achievement (Schmitt, 1990; Taboada, Tonks, Wigfield, & Guthrie, 2009), and numerous studies have documented changes in reading comprehension as a result of strategies instruction (National Institute of Child Health and Human Development, 2000). Strategies that are thought to support comprehension include monitoring and clarifying comprehension difficulties, predicting, questioning, and summarizing (García, Pearson, Taylor, Bauer, & Stahl, 2011). Strategies such as these are thought to play an important role in the way readers construct meaning from text and the way they monitor and repair their meaning construction when comprehension challenges arise. Previous correlational research has shown that facility with these higher level strategic processes makes an independent contribution to reading comprehension after controlling for word reading and verbal ability (Cain, Oakhill, & Bryant, 2004).

Although much of the research on strategy-centric reading comprehension instruction has been conducted in monolingual settings, strategies are also thought to support the reading comprehension of emergent bilinguals (Olson & Land, 2007). For example, in a recent investigation of the relationship between teachers’ instructional practices and elementary students’ growth in comprehension, Silverman and colleagues (2014) found that for monolingual students, frequency of instructional attention to strategies was not associated with comprehension growth but was related to positive gains for bilingual students. However, the current emphasis on cognitive and metacognitive strategies in reading comprehension instruction has been called into question by scholars who argued that readers can benefit from lessons in which they interact directly with the content in the text without the need for overt strategies (McKeown, Beck, & Blake, 2009; Wilkinson & Son, 2011). Reviews of the research on strategies instruction (e.g., Davis, 2010; Rosenshine & Meister, 1994) have shown that the effectiveness of strategy-centric comprehension instruction varies tremendously across studies, with some studies showing small or null effects on comprehension achievement. Thus, there is a need for better understanding of the role of strategic expertise in predicting reading comprehension relative to other hypothesized correlates of comprehension.

Language Proficiency as a Correlate of Reading Comprehension

Research examining the role of language proficiency has revealed its robust effect on first- and second-language (L2) reading (Apel, Wilson-Fowler, Brimo, & Perrin, 2012; Yamashita & Shiotsu, 2015). Comparisons of the relationship between language proficiency and reading have suggested that there may be different relationships between the two for bilingual (in their L2) and monolingual readers (Butler & Hakuta, 2009; Saunders & O’Brien, 2006). This possibility motivated our interest to examine English-language proficiency in this study in a sample of students who vary in their English-language backgrounds and proficiency levels.

Previous research has focused on several facets of language proficiency, including vocabulary, morphology, and syntax. Vocabulary knowledge has been consistently found to predict reading comprehension (Liebfreund, 2015; Silverman et al., 2015). For example, Droop and Verhoeven (2003) found that reading comprehension in fourth-grade monolingual and emergent...
bilingual students was influenced by vocabulary and morphosyntactic knowledge. Vocabulary knowledge was a particularly strong predictor of reading comprehension for emergent bilinguals in the sample who were still developing in their L2. In another study, Proctor, Carlo, August, and Snow (2005) found that English–Spanish bilingual students’ L2 vocabulary was directly and indirectly related to L2 reading comprehension, with the indirect relationship mediated by L2 listening comprehension. In a more recent cross-sectional analysis of elementary school students, researchers also found that vocabulary indirectly affected reading comprehension through its effect on both word recognition and listening comprehension (Language and Reading Research Consortium, 2015). Morphological awareness has also been found to relate to reading comprehension for both monolingual and bilingual students. For example, Carlisle (2000) found that morphological awareness was positively related to reading comprehension in fifth-grade monolingual students. In that study, morphological awareness was measured using a task in which readers completed sentences using morphologically derived forms of words they were given. Kieffer and Lesaux (2008) also examined the relationship of morphology and reading comprehension among ELs using a similar measure and found a robust relationship for fifth-grade students even after controlling for vocabulary, word reading, and phonological awareness. In both studies, the researchers found that the relationship of morphology and comprehension was stronger for older elementary students, suggesting a developmental trend in the way monolingual readers and ELs benefit from their increasing morphological knowledge as they move into the middle school grades.

Finally, knowledge of syntax (or grammatical knowledge at the sentence level), although less studied, may also be related to reading comprehension (Silverman et al., 2015). In a study of adult L2 readers, Shiotsu and Weir (2007) found that syntactic knowledge in the L2 made a significant contribution to L2 reading above and beyond the contribution of vocabulary. This finding was obtained across three separate analyses with L2 readers at different levels of L2 proficiency and experience. Furthermore, recent scholarship has suggested that sentence-level grammatical patterns used in academic texts can impact meanings that are made available to readers and, by extension, the way readers construct meanings from texts within a discipline (Frantz, Starr, & Bailey, 2015).

**Content Knowledge and Reading Comprehension**

Readers’ prior knowledge has been theorized as playing a strong role in how meaning is constructed with a text (Aukerman, Brown, Mokhtari, Valencia, & Palincsar, 2015; Duke et al., 2011). From a schema-theoretic perspective, prior knowledge is thought to provide an ideational scaffold for the reader, resulting in superior comprehension and learning of information that fits the knowledge structures already in place (Anderson, Spiro, & Anderson, 1978). Similarly, in Kintsch’s (2004) construction–integration model, prior knowledge is needed to supplement information provided by the text in order for the reader to construct a situation model—a mental representation of the propositions in a text integrated with outside knowledge.

Prior knowledge is thought to be especially important in the comprehension of expository science texts (McNamara, Floyd, Best, & Louwerse, 2004). Researchers have found that prior knowledge is a strong predictor of adult readers’ comprehension of complex science texts even after controlling for overall reading ability (Ozuru, Dempsey, & McNamara, 2009). Prior knowledge is also considered to be a predictor of L2 comprehension (Chen & Donin, 1997). However, there are questions about the importance of prior knowledge in explaining L2 comprehension relative to other predictors (McNeil, 2011). For example, Usó-Juan (2006) examined the relationships of English-language proficiency, discipline-specific knowledge, and English reading comprehension in a sample of 380 university students in Spain who were learning English as a foreign language. Participants read six expository English texts, excerpted from content textbooks in several different disciplines. Using regression models, the researcher found that discipline-specific knowledge explained between 21% and 31% of the variance in L2 reading comprehension, whereas language proficiency explained approximately twice as much of the variance. These findings suggest that more research is needed to understand the relative contributions of predictors of reading comprehension that have not traditionally been examined simultaneously in linguistically diverse populations.

**Epistemic Thinking in Disciplinary Reading**

The present study was influenced by recent scholarship that has foregrounded the importance of discipline knowledge—understanding the way knowledge is created, disseminated, and revised within a disciplinary tradition—in how readers make sense of and learn from text (Moje, Stockdill, Kim, & Kim, 2011; Shanahan, Shanahan, & Misischia, 2011). Because texts in science serve to codify the findings of past investigations and to provide directions for future inquiry (Greenleaf et al., 2011), these texts call on readers to engage in discipline-specific forms of epistemic thinking related to their “knowledge of knowledge-creation practices within a
A relationship between epistemic beliefs and reading comprehension has been proposed in research with both adult and adolescent readers, but the specific nature of this relationship is still not fully understood. In a study of college students who had differing levels of sophistication in their epistemological beliefs, Bråten and Strømsø (2006) asked participants to read seven texts written from multiple perspectives. The researchers found that readers who believed that knowledge is evolving and socially constructed were able to integrate information across multiple texts, whereas those with so-called naive epistemological beliefs (i.e., beliefs that knowledge is objectively true and given from a higher authority) benefited from reading the information in a textbook format.

In a study of sixth-grade students’ processes of multiple-text comprehension, Barzilai and Zohar (2012) also categorized students based on their epistemic beliefs and found that evaluativists (who viewed knowledge as individually constructed judgments) outperformed absolutists (who viewed knowledge as objectively certain) in their identification of authors’ points of view, frequency of comparisons across different texts, and their use of multiple sources to support their arguments, but they did not differ significantly in their use of source evaluation strategies. This work suggests that although prior knowledge and strategic processes have been heavily studied as predictors of comprehension, personal epistemology could potentially be an important factor as well.

Multiple-Text Comprehension

To maximize their content learning in science and other disciplines, readers must be able to integrate complementary and conflicting information across multiple sources (Stadtler & Bromme, 2013) and attend to the varying levels of authorial credibility and epistemic uncertainty of the knowledge represented across texts (Bråten, Britt, Strømsø, & Rouet, 2011). These two processes (intertextual integration and credibility monitoring) are central features of the documents model (Bråten et al., 2011; Perfetti, Rouet, & Britt, 1999). The documents model extends Kintsch’s (2004) construction-integration model of single-text comprehension and has informed much of the recent work on multiple-text comprehension. This work highlights the need for readers to attend to textual metadata—information about the authors, their credentials, and potential biases and areas of expertise—to form an integrated mental model of the issue discussed across multiple texts. This process requires a more complex strategic repertoire than what would typically be used in single-text reading. In a review of the literature on strategies for reading multiple texts, Afflerbach and Cho (2009) organized these strategies into three categories: strategies that facilitate identifying and learning new information, monitoring one’s meaning construction process, and evaluating texts and authors in light of the reader’s goals.

Research conducted with adult readers has contributed to the field’s understanding of multiple-text comprehension processes. For example, in a study of how adult readers interact with multiple online sources (Goldman, Braasch, Wiley, Graesser, & Brodowinska, 2012), undergraduates were identified as either better or poorer learners based on their gains on a conceptual assessment before and after reading seven websites. The researchers found that better learners were more aware of the credibility of website information and were more likely than poorer learners to visit reliable websites. Quantified think-aloud data showed that better learners were also more strategic than poorer learners, as evidenced by the better learners’ more frequent use of monitoring and evaluation strategies, which informed their navigation patterns as they moved through the texts.

In one of the few studies that examined the processes of multiple-text comprehension with young adolescent readers, Wolfe and Goldman (2005) asked sixth graders to think aloud while reading conflicting historical accounts. The researchers found that the participants engaged in forms of intertextual processing characteristic of expert historical reading, particularly reasoning about an event using multiple conflicting interpretations of what may have actually happened. The researchers also found that students’ scores on a measure of text-based historical reasoning were positively related to their frequency of self-explanations (linking information across texts or to prior knowledge) and their frequency of connecting information within a text.

The Present Study

As reviewed above, previous research has supported the conclusion that strategic expertise, language proficiency, content knowledge, and epistemological beliefs are related to reading comprehension. These predictors align with a socioconstructivist view of reading and learning from texts in the disciplines. This view theorizes that readers draw on their strategic expertise, linguistic resources, and prior content knowledge and beliefs to construct situated understandings that fit the meanings offered by the author/text and the readers’ expectations and purposes for reading (Duke et al., 2011; RAND Reading Study Group, 2002). However, questions remain about the relative contributions of these
factors in the way readers construct understandings of disciplinary texts, particularly among linguistically diverse readers and in the context of multiple-text reading.

For instance, our review of the literature leads us to question the relative importance of traditional single-text strategy expertise. It may be that strategy expertise plays only a minor role in explaining differences in expository comprehension compared with prior knowledge and English-language proficiency, particularly among linguistically diverse readers. Additionally, our review suggests that the role of epistemic beliefs in explaining comprehension differences is still largely unknown. Although epistemology is gaining attention as a possible factor for explaining variation in comprehension, it has not been heavily studied in linguistically heterogeneous young adolescent populations.

Thus, the theory examined in this study is that although all of these hypothesized predictors may be useful contributors to reading comprehension of multiple expository texts, they are likely not equal contributors. Because few studies have directly examined these predictors simultaneously, the field lacks knowledge of their relative contributions. Given that traditional measures of strategy expertise and epistemology may not be revealing of the intricate meaning-making processes that readers employ in the moment while reading complex texts on science topics, there is also a need in the field for detailed qualitative descriptions of these processes.

Our mixed-methods study addresses these gaps in the field’s knowledge by examining these two questions:

1. What are the relative contributions of strategy expertise, English-language proficiency, content knowledge, and epistemic beliefs to monolingual and bilingual middle school students’ comprehension of multiple expository science texts?

2. What disciplinary meaning-making features are evident in students’ interactions with multiple texts on a scientific topic?

The first question was examined using correlation and regression analyses and the second question using verbal think-aloud data that reveal features of students’ meaning-making processes that could not be captured using traditional quantitative measures.

Methods

Participants

The sample consisted of 83 students in grades 5–7, recruited from three schools that serve culturally and linguistically diverse communities in a large city in the Southwestern United States. At each school, we worked with teachers who identified and invited students to participate. Teachers were asked to invite students who would represent the range of reading achievement levels of students on their campuses and the varying language backgrounds of their school communities. Teachers explained the study to their students using materials provided by the researchers. Students were told that the researchers were conducting a study on reading comprehension and language and that the study would involve the completion of activities and tests with the researchers, including surveys and tests of reading and language ability. Parental consent and student assent were obtained for all participants prior to beginning the study. Participants were given a $15 gift card for their participation.

The sample comprised 39 males and 44 females. The grade distribution was as follows: 26 fifth graders, 22 sixth graders, and 35 seventh graders. Students ranged from 10 to 14 years of age, with a mean age of 12. According to responses provided by students, 18.1% of their parents had earned a high school diploma as their highest level of formal education, 21.7% had completed some college, and 53% had a college or graduate degree. Twenty-three students (27.7%) were monolingual English speakers, and the rest (N = 60) reported either having a different home language or speaking both English and an additional language at home. Twenty-three of the 60 bilingual students were emergent bilinguals who were classified as ELs by their schools at the time of data collection. The remaining 37 were proficient bilinguals who were not considered in need of English-language support by their schools. This group included students who were never identified as ELs and students who had previously been redesignated as non-ELs according to school criteria. The most common home language among the bilingual students was Spanish. Other home languages included Arabic, Turkish, Punjabi, Hindi, and Bengali.

After data collection for the full sample was completed, we purposively selected 10 students from the larger sample to participate in the qualitative think-aloud portion of this study (see Table 1). The sample size of 10 was chosen to minimize the burden of this research on the participating schools and to allow for in-depth analysis of a few students’ reading processes. We chose to only sample bilingual students because of our interest in adding to the field’s knowledge of bilingual students’ reading practices in science, a topic that has not been heavily studied in previous work. From the group of students who identified as bilingual, we purposively selected students who represented the full range of scores on the expository comprehension measure in the larger sample while also ensuring that the
group would include membership in all four quartiles of English-language proficiency in our sample.

The resulting think-aloud sample consisted primarily of proficient bilingual students. Only one of the students (reader E) was labeled by her school as an EL, but eight of the 10 indicated that they had previously received English as a Second Language services at their schools. When asked about the frequency of English use at home with their parents, students responded as follows: One student indicated speaking only English with her parents, one speaks mostly English, six speak some English, and one speaks a little English.

**Procedure**

Students were administered the assessments and surveys at their schools by research personnel in group and individual sessions during the school day. Teachers on each campus assisted us in scheduling the sessions to minimize disruptions to class time. During group sessions, students completed measures of epistemological beliefs, prior knowledge, morphology, expository comprehension, and strategy expertise, in that order. Students were pulled from class for the individual sessions during times when they and the teachers agreed that they could miss typically occurring classes and activities (e.g., homework help time, other flexible times in their schedules). During the individual sessions, students completed measures of vocabulary, productive syntax, and receptive syntax, in that order. Students spent 90–120 minutes completing group-administered assessments (with breaks provided or spread over two separate sessions) and one hour meeting individually with a researcher.

The 10 students selected for the think-aloud analyses met with a researcher for a third time for this activity. These meetings occurred approximately two weeks after the completion of initial data collection. Prior to data collection, all instruments and protocols were piloted with four participants to examine and revise the timing, order, and language demands of the protocols.

**Data Sources**

**Expository Comprehension With Multiple Texts**

This construct was measured using a researcher-designed assessment patterned after sentence verification (SV) and
inference verification (IV) tests used in previous comprehension research (e.g., Bråten, Strømso, & Britt, 2009). The measure comprised 35 dichotomously scored items that students answered after reading two texts about the decision to reclassify Pluto as a dwarf planet (see the Appendix, which is available as supporting information for the online version of this article). Students were not allowed to look back at the text when answering the SV items, but they were allowed to look back when answering the IV items. Items related to both texts were interspersed across the SV and IV sections of the test, along with items that required intertextual integration of both texts. Students completed the untimed assessment in a group session supervised by the researchers.

To ensure construct validity, the 19 SV and 16 IV items were developed using an iterative process of item revision based on expert review and response process data from pilot participants. Drafts of the instrument were reviewed by doctoral students who completed all items and gave feedback on item difficulty, clarity, and construct alignment. This feedback was used to delete and revise items across multiple iterations of the instrument. The penultimate version was field-tested with four children of the same ages as the intended sample. They were observed and audiotaped while completing the instrument in a structured interview format to obtain additional information about response processes. This information was used to further revise items for clarity and construct alignment.

Using an item response theory model (Wilson, 2005), we analyzed item fit statistics (point-biserial correlation with overall score, item characteristic curves, and infit mean squares) to identify and eliminate items with unexpected item behavior. This resulted in the deletion of two items. The scale computed with the remaining 33 items shows reasonable internal consistency (Kuder–Richardson reliability = .73) and concurrent validity (r = .66 with a state-mandated reading test that primarily assesses single-text comprehension but also includes paired text items). Additional validity evidence comes from the pattern of students’ performance across different types of items. As expected, students performed better on the SV items (mean percentage correct = 77.5, standard deviation [SD] = 14.8) than on IV items (mean percentage correct = 66.6, SD = 13.6). The mean score on intertextual items (62.6%) was lower than the mean score on single-text items (75.3%). The SV and IV scales did not map onto separate factors, nor did the intertextual and single-text items, so all items were combined into one score. Ability estimates (in logits) derived from the item response theory model were used in the analyses instead of raw scores.

The texts used on this measure describe the process and controversy of the International Astronomical Union’s (IAU) development of a new definition of planets and dwarf planets, which resulted in Pluto’s reclassification. This topic was chosen to highlight the epistemic uncertainty and fallible nature of scientific inquiry consistent with the perspective that “scientific knowledge is a byproduct of negotiated agreements among people concerning the nature of the world” (Gergen & Dixon-Román, 2014, p. 6). The two texts were written by the researchers (based on Internet sources) in consultation with a science educator. They were designed to emulate typical expository articles that one might find on the Internet and to appear to be written by different people representing different perspectives on the Pluto decision. As shown in Table 2, the texts differed in their degree of transparency of the authorial voice (Paxton, 1999) responsible for the creation of the text. Our own Internet searches for information about Pluto revealed that texts on this topic tended to vary in their agreement with the Pluto decision, with arguments against the decision often coming from non-scientific sources. We wanted our texts to reflect the variations in tone and authorial credentials that we encountered in authentic online texts.

English-Language Proficiency

A multidimensional measure of English-language proficiency was constructed from students’ scores on vocabulary, morphological awareness, productive syntax, and receptive syntax measures. Students were administered the fourth edition of the Peabody Picture Vocabulary Test (PPVT–4; Dunn & Dunn, 2007), a derivational morphology decomposition task used in previous research (Kieffer & Lesaux, 2008), and the Sentence Assembly subtest from the fourth edition of the Clinical Evaluation of Language Fundamentals (CELF–4; Semel, Wiig, & Secord, 2003) as a measure of productive syntax. The derivational morphology measure had a reliability of .87 in the present sample. The PPVT–4 and the CELF–4’s Sentence Assembly subtest are widely used norm-referenced measures of language with published internal consistency reliabilities of .96–.98 and .80–.83, respectively, for the age range of the participants in this study. As a measure of receptive syntax, we administered a researcher-designed grammaticality judgment task, adapted from a previous study with adult bilingual learners (Huang, 2014). Participants listened to a series of 66 sentences and judged each one as grammatically acceptable or unacceptable. The measure was found to exhibit reasonable internal consistency within our sample (Kuder–Richardson reliability = .89).

Scores for these four language measures were all positively correlated with each other and with the expository comprehension outcome. To maintain a parsimonious regression model and to minimize collinearity, the four
scores were reduced to a single principal component score, the sole component extracted in a factor analysis using the four language variables. The correlations of the component score and the individual language scores were as follows: .91 with receptive syntax, .88 with morphology, .87 with vocabulary, and .61 with productive syntax.

### Strategy Knowledge

Students were administered 20 items from the Metacomprehension Strategy Index (MSI; Schmitt, 1990), a multiple-choice assessment that asks students to choose which strategy is most useful for helping them understand what they are reading. The six strategies covered by the measure are predicting and verifying predictions, previewing, purpose setting, self-questioning, using background knowledge, and summarizing to repair comprehension. This instrument is commonly used to measure knowledge of reading strategies and was found to be reliable and valid by the instrument developer (Schmitt, 1990).

Like others who have used this measure (O’Reilly & McNamara, 2007), we modified the items to make them more suitable for examining knowledge of strategies used for expository reading (e.g., replacing the word story with the word text). Students completed the first two sections of items in the three-section survey (“Before Reading” and “During Reading”). The 20-item scale was found to have low internal consistency in this sample (Kuder–Richardson reliability = .61), which led us to examine item-level response patterns. A 10-item scale was found to be optimal in this sample, with an internal consistency reliability of .68. The 10-item scale includes items from all six of the strategy categories measured on the full instrument.

### Strategy Awareness

The Metacognitive Awareness of Reading Strategies Inventory (MARSII; Mokhtari & Reichard, 2002) is a 30-item survey of perceived strategy use that asks students to rate individual strategies on a 5-point numerical scale based on frequency of use when reading academic and school-related texts. The instrument developers reported high reliability coefficients for the instrument in middle school and high school samples (Cronbach’s α between .86 and .93; Mokhtari & Reichard, 2002). As evidence of validity, they reported that MARSII scores are related to students’ self-reported reading ability. The items showed high internal consistency (Cronbach’s α = .86) in this sample. This measure was selected for use in addition to the MSI because we conjectured that these instruments would measure different dimensions of strategic expertise, based on our own interpretations of the different formats of the surveys. The strategy awareness scale (MARSII) appears to measure self-reported generalized frequency of use of strategies, whereas the strategy knowledge measure (MSI) captures readers’ ability to select the strategies that are deemed optimal by the instrument authors, thus reflecting alignment with a particular view of

---

**TABLE 2**

Comparison of the Three Researcher-Designed Texts Used in This Study

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Text 1</th>
<th>Text 2</th>
<th>Text 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>“How a Disappearing Planet Changed My Life”</td>
<td>“Our Smallest Planet Is No Longer a Planet”</td>
<td>“Pluto Huggers Unite! Four Reasons Pluto Should Still Be a Planet”</td>
</tr>
<tr>
<td>Length</td>
<td>677 words</td>
<td>597 words</td>
<td>352 words</td>
</tr>
<tr>
<td>Flesch–Kincaid readability</td>
<td>7.9</td>
<td>8.3</td>
<td>8.2</td>
</tr>
<tr>
<td>Content</td>
<td>Personal account of an astronomer’s childhood affection for Pluto</td>
<td>Textbook-like account of the IAU’s decision and other relevant definitional information about objects in space</td>
<td>Overt but not entirely compelling argument for reclaiming Pluto’s planethood</td>
</tr>
<tr>
<td>Transparency of author’s identity</td>
<td>Transparent author, who is named and identified as an astronomer in the first line of the text; first-person writing</td>
<td>Hidden author, not named; third-person writing with author’s identity suppressed</td>
<td>Semitransparent author, plural first-person writing (“we”); the credentials or specific identities of the “Pluto huggers” not given</td>
</tr>
<tr>
<td>Stance</td>
<td>Subtle opposition to the IAU’s decision but softened by a respect for scientific inquiry</td>
<td>Implied support of the IAU’s decision embedded within “objective” recounting of the process</td>
<td>Overtly opposed to the IAU’s decision; in favor of Pluto’s status as a planet</td>
</tr>
<tr>
<td>Tone</td>
<td>Personalized, passionate, but not dogmatic</td>
<td>Relatively objective and distant</td>
<td>Dogmatic and impassioned</td>
</tr>
</tbody>
</table>

Note. IAU = International Astronomical Union.
strategy effectiveness beyond self-reported frequency of use.

**Content Knowledge**

Content knowledge was assessed using a researcher-designed measure of prior knowledge of Pluto, planets, and astronomy. This measure was administered to students prior to completing the expository comprehension measure. Students were presented with a series of statements and asked to judge them as accurate or inaccurate based on their prior knowledge (e.g., "Pluto was discovered in the 1990s" [inaccurate]; "Neptune is sometimes farther from the sun than Pluto" [accurate]). The measure entailed 14 dichotomously scored items. Item analyses were conducted to verify that all items showed a positive point-biserial correlation with overall score on the measure. These items were reviewed by experts and revised for clarity and construct alignment over several iterations to ensure validity.

**Epistemic Beliefs**

Items from the Epistemological Belief Questionnaire (Schommer-Aikins, Mau, Brookhart, & Hutter, 2000) were used to measure students’ beliefs related to knowledge. Two subscales that were validated in previous work with adult learners (Schommer, 1990) and subsequently tested with middle school students (Schommer-Aikins et al., 2000) were selected for this study: stability of knowledge (four items) and structure of knowledge (nine items). Students were presented the items and asked to rate their agreement on a 5-point scale. Items were scored so higher scores indicated stronger alignment with a view of knowledge as complex and dynamic. Consistent with previous work that has shown epistemological beliefs to be multidimensional, the 13 items did not show strong internal consistency when treated as a single score, prompting us to examine the two factors separately. Two items were retained in the stability-of-knowledge scale and four items in the structure-of-knowledge scale.

The items used in this analysis in the stability-of-knowledge scale were: "Scientists can get to the truth if they just keep searching for it," and "If scientists try hard enough, they can find the truth to almost everything." These two items were strongly correlated (Cronbach’s $\alpha = .73$). During analysis, the scale was reversed. A student’s mean agreement with these two items was used as an indication of where he or she fell on a continuum from viewing knowledge as static and objective (low score, or high agreement) to viewing knowledge as dynamic and evolving (high score, or low agreement).

The four items used for the structure-of-knowledge scale were: "I can depend on facts written in my school books for the rest of my life," 2 "It is hard to learn from a textbook unless you start at the beginning and learn one chapter at a time," 3 "You will get mixed up if you try to combine new ideas in a textbook with what you already know," and "Being a good student generally involves memorizing facts" (Schommer-Aikins et al., 2000, p. 123). Participants’ agreement with these items was used as an indicator of their beliefs about how knowledge is constructed, with one end of the continuum (low scores) representing knowledge as discrete pieces of information and the other end (high scores) representing knowledge as “complex, interrelated concepts” (Schommer-Aikins et al., 2000, p. 121). These four items exhibited only marginal internal consistency (Cronbach’s $\alpha = .47$).

This scale was retained in the analysis because it was found in preliminary analyses to correlate with expository comprehension more strongly than the stability-of-knowledge scale, and further exploration of this relationship was warranted. Our findings corroborate previous work in which it was found that beliefs related to the structure of knowledge are difficult to reliably measure in middle school students whose views of knowledge are still developing (Schommer-Aikins et al., 2000).

**Think-Aloud Protocol**

Ten students sampled for the think-aloud portion of this study met individually with a trained researcher who administered a 45-minute structured think-aloud protocol. Students were told that the researchers wanted to learn more about their thinking while reading multiple texts. The researcher explained the think-aloud process and gave the students instructions to comment on whatever they were thinking about while reading the texts aloud. Students completed a practice paragraph, with researcher guidance and feedback, to gain familiarity with the think-aloud process. Then, students read three texts and verbalized their thinking spontaneously and at preplanned stopping points in response to researcher prompts. The three texts used for the think-aloud were the two texts about Pluto’s reclassification that the participants had previously read on the expository comprehension measure and a third researcher-created text on the same topic (see Table 2).

We built in multiple open-ended prompts throughout the texts to maximize the amount of verbal data generated. In addition to regular reminders to think aloud whenever they had something to share, we stopped participants at 16 preplanned points and gave these prompts: “Tell me what you are noticing or thinking so far about this text”; “Is there anything else you are noticing, wondering, or asking yourself?”; “Are you noticing any connections or relationships between this text and the other texts?”; and “What is challenging or
confusing for you?” At the end of the three texts, participants were asked to respond to standardized postreading questions (available in the Appendix). Verbal data were audiotaped and transcribed.

Quantitative Analyses

Correlation and hierarchical multiple regression analyses were conducted to examine the relative magnitude of association of hypothesized predictors with expository reading comprehension. For all analyses involving decisions about statistical significance, we used the conventional Type I error rate (α) of .05. The selection of hypothesized predictors was based on prior research reviewed in a previous section of this article. In addition to these substantive predictors, we also considered a panel of demographic control variables that we conjectured might be related to reading comprehension: students’ EL status as determined by school district criteria, highest education of a student’s parent (as a proxy for socioeconomic status), gender, age, and whether the student self-identified as bilingual.

To maintain the parsimonious regression model necessitated by the modest sample size of this study, we chose to include all of the substantive predictors in the model but to be selective in the inclusion of control variables. Control variables were screened for their relationships with the outcome prior to inclusion in the model. EL status was the only demographic variable that was found to correlate with expository comprehension and contribute unique variance in a regression model including the other potential controls. The final regression model was constructed using a hierarchical procedure, with predictors entered in the following order: EL status, English-language proficiency, prior content knowledge, two strategy scales, and two epistemology scales.

Using guidance provided by Cohen, Cohen, West, and Aiken (2013), we verified that the final model meets all of the assumptions of regression. Scatterplots were examined for each bivariate relationship with the outcome to rule out the possibility of nonlinear relationships. We also verified the linear form of the relationships through visual inspection of the plots of each predictor to the residuals and of the predicted values to the residuals. We also examined the above plots to visually detect any increased or decreased variance at different levels of the independent variables. No patterns indicative of heteroscedasticity were detected. We inspected the histogram and the normal quantile-quantile plot of the residuals to verify their normality. Finally, we examined the tolerance value for each predictor as an indication of the proportion of variance in each predictor that is nonoverlapping with other predictors. Tolerances range between .39 and .89, with corresponding variance inflation factor values between 2.54 and 1.21. These values are well within acceptable values of collinearity.

Qualitative Analysis of Think-Aloud Data

Previous research using quantified verbal data has made substantial contributions to the field’s understanding of reading processes (e.g., Goldman et al., 2012; Wyatt et al., 1993). In this study, we chose to employ a constructivist qualitative approach instead, using constant comparative analysis (CCA; Glaser, 1965; Strauss & Corbin, 1994) to construct a description of the salient patterns in the participants’ reading processes rather than quantifying their verbal behaviors. Coding proceeded according to a recursive three-phase process: open coding and memoing to identify the salient categories present in the data, axial coding to further refine the categories and link ideas across and within categories, and selective coding to centralize and superordinate the codes into core categories that represent the most defensible interpretations of the data (Strauss & Corbin, 1998).

We acknowledge that our CCA was different in some ways from what some would consider purer forms of grounded theory building often associated with this form of analysis. For instance, our CCA was both inductive and theory driven. We remained open to unplanned categories in the data, but our coding was ultimately driven by our purpose of understanding middle school students’ uses of discipline-specific reading practices across multiple texts. Also, given the intent of our analyses, our findings rely more on description than on the development and revision of theory. Despite these needed deviations from more grounded approaches to qualitative analysis, we found CCA to be the most appropriate way to fully characterize the concepts and relationships evident in the data. We describe our analytic process chronologically in the following sections, although in practice, it was more recursive than suggested by this chronology.

First, during open coding, we noticed patterns in students’ processes that led us to begin marking three features of students’ verbal data: strategic function, focus of attention, and what the think-alouds revealed about the concepts represented in the reader’s mental representations of the texts. These initial codes, which took shape through repeated reading of the transcripts, were informed by the codes described by Goldman and colleagues (2012) in their analysis of undergraduate students’ think-aloud behaviors. Strategic functions that were evident in the data included paraphrasing or repeating text information, self-explanation (i.e., reasoning, interpretation beyond the level of paraphrasing or
repetition), mental monitoring or questioning, evaluation or credibility monitoring, and intertextual linking. Attention foci were defined as the source to which the student applied the strategic process, including attention to the authorial presence behind the text, the content of the text, the disciplinary knowledge related to knowledge creation implied in the text, and the language/structure of the text. To examine the concepts that readers were attending to in their construction of meaning from the text, we coded for central ideas discussed in the texts (e.g., the IAU’s decision-making process, reasons for Pluto’s reclassification, the New Horizons mission).

This first pass through the data (open coding) gave us a framework for interpreting the way students strategically interacted with the multiple texts and the mental representations that they constructed. Then, as axial coding proceeded, we reached consensus around three broad concepts that best characterize the patterns and variations in the ways that readers constructed their understandings of the texts: working toward coherent situational understanding of the topic through the use of self-explanation and monitoring, integrating information across multiple texts by theorizing identities for the authors and identifying intertextual linkages, and indicators of epistemic thinking related to the discipline of science. These three superordinate categories emerged from the initial coding categories described previously and were developed through discussion and repeated reading of the raw data.

Finally, we engaged in selective coding to help us more clearly understand the emerging concepts. We did this by developing case summaries of each student related to the specific features of students’ reading processes suggested by these concepts (self-explanation, monitoring, intertextual linkages, theorizing authorial identities, and epistemic indicators). The summaries were constructed by repeatedly returning to the data and asking the question, How is this feature of multiple-text reading playing out in this student’s textual interactions? To finalize the analyses, we identified the consistencies and variations across cases. This was done by using the case summaries and the raw data as we asked, What is true of all or most cases in the think-aloud sample? and What is unique about each case that might reveal important variations in multiple-text reading worth exploring in future work? These are the consistencies and variations that we report in the article.

The intended endpoint of our CCA was a description of students’ meaning-making processes. To maximize the trustworthiness of this description, credibility was established through triangulation of findings across multiple participants and through consensus building among the research team during data analysis (Denzin & Lincoln, 2008). At the end of data analysis, to perform a confirmability check, a meeting was held between Davis (first author) and a research assistant who was not involved in the initial analyses. We traced the findings back to the raw data and selected examples for reporting that best exemplified the data set. We report examples of raw data throughout the Findings section to facilitate readers’ assessments of the potential transferability of these findings to other settings.

Findings

Descriptive statistics for all measures used in this analysis are shown in Table 3.

Participants varied in their performance on the expository comprehension measure, with scores ranging from 39% to 100% correct. Their age-normed standard scores on the PPVT–4 ranged from 35 (<0.01 percentile) to 127 (96th percentile), with a mean score near the 50th percentile according to U.S. national norms (Dunn & Dunn, 2007). According to the age norms for the productive syntax measure (Semel et al., 2003), our sample mean of 9.16 corresponds approximately to an age equivalent of 9 years 9 months, and scores on this measure span the age equivalents of 7 years 0 months to 17 years 11 months. National norms for the other language measures are not available, but we interpret the vocabulary and productive syntax scores as evidence that the sample included students with a wide range of English-language proficiency levels.

Scores on the content knowledge measure indicated a wide range of prior knowledge of the topic. Results of the strategy awareness survey (MARS) indicate that students tended to fall in the middle of the 5-point scale in their self-reported strategy use. Students’ scores varied tremendously for strategy knowledge, with a mean of about 60%. The mean score of about 2 (out of 5) on the stability-of-knowledge scale indicates that students tended to align with a view of knowledge as unchanging and objective. On the structure-of-knowledge scale, the mean score of approximately 3 indicates that students tended to be in the middle of the scale (i.e., neither agree nor disagree) when asked to respond to statements about the simplicity or complexity of how knowledge is structured.

Correlations and Multiple Regression Results

Bivariate correlations among all variables used in the analyses are shown in Table 4. The variables found to be statistically significantly related to students’ scores on the expository comprehension measure were, in order of magnitude, English-language proficiency ($r = .56$), prior knowledge ($r = .42$), structure of knowledge ($r = .27$), and strategy knowledge ($r = .23$). Additionally,
EL status was also found to be related to expository comprehension \( (r = –.36) \), with students identified by their schools as ELs scoring lower on average on the measure. Strategy awareness and strategy knowledge were only weakly correlated \( (r = .29) \), and strategy awareness was not related to the comprehension outcome \( (r = .05) \), supporting our conjecture that these surveys measure different dimensions of strategic expertise.

Table 5 shows the regression coefficients for the demographic control variables that were screened for inclusion in the multiple regression analysis. The results of the hierarchical multiple regression analyses, with expository comprehension as the dependent variable and the hypothesized predictors added in five separate steps, are in Table 6. In model 1, EL status, the one control variable that was found to be related to expository comprehension, was associated with 12.6% of the variance in the outcome when entered alone. When language proficiency was added (model 2), it was found to account for an additional 19.3% of the variance in expository comprehension. Prior content knowledge (model 3) was responsible for an added 4.6% of explained variance beyond the contributions of EL status and language proficiency. The two strategy measures (model 4), however, only added a small and nonsignificant improvement in \( R^2 \) (4.4%). The epistemology scales (model 5) were also found to be associated with only a small increase in the model \( R^2 \) (3.6%). The final model explained 44.5% of the variance in expository comprehension scores.

Comparing the magnitude of the standardized coefficients for the hypothesized predictors in the full model (model 5), we found that English-language proficiency was the strongest predictor of expository comprehension \( (\beta = 0.476) \), followed by prior content knowledge \( (\beta = 0.242) \). Stability-of-knowledge beliefs and strategy knowledge had weak relationships with the outcome and only approached statistical significance \( (\beta = –0.178^3 \text{ and } \beta = 0.175, \text{ respectively}) \). After accounting for the other relationships, structure-of-knowledge beliefs was no longer a significant predictor of expository comprehension \( (\beta = 0.120) \), and neither was strategy awareness \( (\beta = 0.044) \). The \( R^2 \) change values for each step in the hierarchical regression (in the final column of Table 6) show that language was uniquely associated with 8.9% of the variance in expository comprehension compared with 4.5%, 3.5%, and 3.6% for content knowledge, strategy, and epistemology, respectively.

### Table 3
Descriptive Statistics \((N = 83)\)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expository comprehension</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Raw score(^a)</td>
<td>72.62</td>
<td>13.73</td>
<td>39.39</td>
<td>100</td>
</tr>
<tr>
<td>• Logit score</td>
<td>1.23</td>
<td>0.62</td>
<td>–0.12</td>
<td>2.71</td>
</tr>
<tr>
<td>Receptive syntax(^a) (grammaticality judgment)</td>
<td>84.64</td>
<td>12.06</td>
<td>49.21</td>
<td>98.41</td>
</tr>
<tr>
<td>Vocabulary (standard score on the Peabody Picture Vocabulary Test, 4th edition(^b))</td>
<td>97.23</td>
<td>18.97</td>
<td>35</td>
<td>127</td>
</tr>
<tr>
<td>Morphology(^a)</td>
<td>85.64</td>
<td>15.85</td>
<td>28</td>
<td>100</td>
</tr>
<tr>
<td>Productive syntax (sentence assembly)</td>
<td>9.16</td>
<td>4.61</td>
<td>1</td>
<td>18</td>
</tr>
<tr>
<td>English-language proficiency (language composite score)</td>
<td>0.00</td>
<td>1.00</td>
<td>–2.37</td>
<td>1.51</td>
</tr>
<tr>
<td>Epistemology subscales</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Stability of knowledge(^c)</td>
<td>1.99</td>
<td>1.00</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>• Structure of knowledge(^c)</td>
<td>2.92</td>
<td>0.81</td>
<td>1.25</td>
<td>4.75</td>
</tr>
<tr>
<td>Prior content knowledge(^a)</td>
<td>70.22</td>
<td>15.49</td>
<td>35.71</td>
<td>100</td>
</tr>
<tr>
<td>Strategy awareness (Metacognitive Awareness of Reading Strategies Inventory(^d))</td>
<td>3.28</td>
<td>0.54</td>
<td>1.60</td>
<td>4.63</td>
</tr>
<tr>
<td>Strategy knowledge (Metacomprehension Strategy Index(^e))</td>
<td>59.64</td>
<td>23.71</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

TABLE 4
Correlation Matrix (N = 83)

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Expository comprehension</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. English learner status^</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Age in months</td>
<td>.166</td>
<td>−.297</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Highest education of parents</td>
<td>.063</td>
<td>−.035</td>
<td>.119</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Bilingual status^</td>
<td>−.030</td>
<td>.383</td>
<td>−.169</td>
<td>.092</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Gender^</td>
<td>−.065</td>
<td>.097</td>
<td>−.274</td>
<td>−.032</td>
<td>.172</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Receptive syntax</td>
<td>.473</td>
<td>−.706</td>
<td>.194</td>
<td>.122</td>
<td>−.199</td>
<td>−.047</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Vocabulary</td>
<td>.564</td>
<td>−.663</td>
<td>.152</td>
<td>.178</td>
<td>−.240</td>
<td>−.052</td>
<td>.786</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Productive syntax</td>
<td>.246</td>
<td>−.198</td>
<td>.112</td>
<td>.142</td>
<td>.056</td>
<td>.111</td>
<td>.404</td>
<td>.336</td>
<td>.446</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. Strategy knowledge</td>
<td>.231</td>
<td>.101</td>
<td>−.100</td>
<td>.050</td>
<td>.242</td>
<td>.016</td>
<td>−.022</td>
<td>−.047</td>
<td>.026</td>
<td>.037</td>
<td>−.006</td>
<td>−.090</td>
<td>−.059</td>
<td>.132</td>
<td>.287</td>
<td>—</td>
</tr>
</tbody>
</table>

^0 was entered for students who were not classified as English learners at the time of data collection and 1 for those who were. ^As reported by students on a survey about their home languages, 0 was entered for students who speak English only at home and 1 for students who indicated speaking another language other than English at home. ^Female = 0, and male = 1. ^Language composite score for receptive syntax, vocabulary, morphology, and productive syntax.

^p < .05. **p < .01.
### TABLE 5
Regression Coefficients for Potential Control Variables Entered Simultaneously as Predictors of Expository Reading Comprehension ($N = 83$)

<table>
<thead>
<tr>
<th>Variable</th>
<th>$B$ (standard error)</th>
<th>$\beta$</th>
<th>$t$</th>
<th>$p$</th>
<th>$\Delta R^2$ (when variable is entered last)</th>
</tr>
</thead>
<tbody>
<tr>
<td>English learner status</td>
<td>-0.497 (0.169)</td>
<td>-0.352</td>
<td>-2.942</td>
<td>.004</td>
<td>.102**</td>
</tr>
<tr>
<td>Age (in months)</td>
<td>0.003 (0.006)</td>
<td>0.056</td>
<td>0.484</td>
<td>.630</td>
<td>.003</td>
</tr>
<tr>
<td>Highest education of parents</td>
<td>0.018 (0.065)</td>
<td>0.031</td>
<td>0.278</td>
<td>.782</td>
<td>.001</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.047 (0.139)</td>
<td>-0.038</td>
<td>-0.336</td>
<td>.738</td>
<td>.001</td>
</tr>
<tr>
<td>Bilingual status</td>
<td>0.179 (0.161)</td>
<td>0.132</td>
<td>1.109</td>
<td>.271</td>
<td>.015</td>
</tr>
</tbody>
</table>

Note. $R^2$ for the full model is .124 (adjusted $R^2 = .065$).

*p < .05. **p < .01.

### TABLE 6
Hierarchical Regression Coefficients for Analysis Predicting Expository Reading Comprehension ($N = 83$)

<table>
<thead>
<tr>
<th>Variable</th>
<th>$B$ (standard error)</th>
<th>$\beta$</th>
<th>$t$</th>
<th>$p$</th>
<th>$\Delta R^2$ (when step is added last)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1 ($R^2 = .126$)</td>
<td>1.361</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Intercept</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• English learner status</td>
<td>-0.488 (0.143)</td>
<td>-0.355</td>
<td>-3.414</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>Model 2 ($R^2 = .318; \Delta R^2 = .193$)</td>
<td>1.197</td>
<td></td>
<td></td>
<td></td>
<td>.089</td>
</tr>
<tr>
<td>• Intercept</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• English learner status</td>
<td>0.102 (0.177)</td>
<td>0.074</td>
<td>0.575</td>
<td>.567</td>
<td></td>
</tr>
<tr>
<td>• English-language proficiency</td>
<td>0.380 (0.080)</td>
<td>0.614</td>
<td>4.753</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>Model 3 ($R^2 = .365; \Delta R^2 = .046$)</td>
<td>0.528</td>
<td></td>
<td></td>
<td></td>
<td>.045</td>
</tr>
<tr>
<td>• Intercept</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• English learner status</td>
<td>0.130 (0.173)</td>
<td>0.095</td>
<td>0.752</td>
<td>.454</td>
<td></td>
</tr>
<tr>
<td>• English-language proficiency</td>
<td>0.330 (0.080)</td>
<td>0.533</td>
<td>4.101</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>• Prior content knowledge</td>
<td>0.009 (0.004)</td>
<td>0.236</td>
<td>2.398</td>
<td>.019</td>
<td></td>
</tr>
<tr>
<td>Model 4 ($R^2 = .409; \Delta R^2 = .044$)</td>
<td>0.048</td>
<td></td>
<td></td>
<td></td>
<td>.035</td>
</tr>
<tr>
<td>• Intercept</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• English learner status</td>
<td>0.071 (0.171)</td>
<td>0.051</td>
<td>0.414</td>
<td>.680</td>
<td></td>
</tr>
<tr>
<td>• English-language proficiency</td>
<td>0.324 (0.079)</td>
<td>0.524</td>
<td>4.118</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>• Prior content knowledge</td>
<td>0.008 (0.004)</td>
<td>0.209</td>
<td>2.137</td>
<td>.036</td>
<td></td>
</tr>
<tr>
<td>• Strategy awareness</td>
<td>0.089 (0.108)</td>
<td>0.076</td>
<td>0.825</td>
<td>.412</td>
<td></td>
</tr>
<tr>
<td>• Strategy knowledge</td>
<td>0.005 (0.002)</td>
<td>0.179</td>
<td>1.914</td>
<td>.059</td>
<td></td>
</tr>
<tr>
<td>Model 5 ($R^2 = .445; \Delta R^2 = .036$)</td>
<td>0.036</td>
<td></td>
<td></td>
<td></td>
<td>.036</td>
</tr>
<tr>
<td>• Intercept</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• English learner status</td>
<td>0.069 (0.168)</td>
<td>0.050</td>
<td>0.409</td>
<td>.684</td>
<td></td>
</tr>
<tr>
<td>• English-language proficiency</td>
<td>0.295 (0.085)</td>
<td>0.476</td>
<td>3.467</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>• Prior content knowledge</td>
<td>0.010 (0.004)</td>
<td>0.242</td>
<td>2.468</td>
<td>.016</td>
<td></td>
</tr>
<tr>
<td>• Strategy awareness</td>
<td>0.051 (0.111)</td>
<td>0.044</td>
<td>0.463</td>
<td>.645</td>
<td></td>
</tr>
<tr>
<td>• Strategy knowledge</td>
<td>0.005 (0.002)</td>
<td>0.175</td>
<td>1.904</td>
<td>.061</td>
<td></td>
</tr>
<tr>
<td>• Stability of knowledge</td>
<td>-0.110 (0.056)</td>
<td>-0.178</td>
<td>-1.955</td>
<td>.054</td>
<td></td>
</tr>
<tr>
<td>• Structure of knowledge</td>
<td>0.092 (0.076)</td>
<td>0.120</td>
<td>1.209</td>
<td>.230</td>
<td></td>
</tr>
</tbody>
</table>

*p < .05. **p < .01.
Analysis of Relationships by Language Group

Table 7 shows the relationships between expository comprehension and each of the hypothesized predictors separately for three subsamples: monolingual English speakers, proficient bilingual learners who were not classified as ELs by their schools, and emergent bilinguals who were classified as ELs. These descriptive analyses were pursued to better understand the role of English-language proficiency in predicting reading comprehension for students who were identified as varying in their language-learning histories and current English-language proficiency. Prior to conducting these analyses, we examined patterns in English-language proficiency across these three groups. Monolingual and proficient bilingual readers did not differ significantly in their English-language proficiency (monolinguals: mean $M = .32$, $SD = .61$; proficient bilinguals: $M = .50$, $SD = .61$; $t = 1.166; p = .248$), but emergent bilinguals ($M = −1.12$, $SD = .95$) scored significantly lower than proficient bilinguals ($t = 8.009; p < .001$).

Squared semipartial correlations, which controlled for the relationship of other predictors to expository comprehension, indicated that for monolingual English speakers in our sample, a larger share of the variance in expository comprehension was accounted for by prior content knowledge (7.2%) and strategy knowledge (6.3%) than by English-language proficiency (0.2%). For bilingual students in the sample (both proficient and emergent), English-language proficiency explained a larger share of the variance (24.0% and 19.5%, respectively) compared with the other predictors.

Qualitative Patterns in Expository Comprehension of Multiple Texts

In the sections that follow, we describe the patterns of typical comprehension processes within the three major categories that emerged in our analyses: how readers used self-explanation and monitoring/questioning to work toward coherent situational understandings, the role of intertextual linkages and theorization of authorial identities in their integration of information across texts, and how readers engaged with the epistemological complexity of scientific knowledge creation described in the three texts.

Table 8 provides case summaries of selected readers’ uses of the five features of expository comprehension subsumed within these three categories. We selected readers A, F, and H to report in this article because they best exemplified the patterned ways of interacting with the multiple texts in this sample. Reader J represents an atypical case who exemplifies the pattern of reader–text interactions that might be considered optimal for constructing an integrative understanding of the three texts.

General Cross-Case Consistencies

Three general consistencies identified across the 10 cases help contextualize the patterns described in the upcoming sections. First, we made a distinction between strategic competence, defined as the ability to successfully use a strategy, and strategic awareness, or knowing the appropriate time and place to deliberately invoke a specific strategy (Andreassen, & Bråten, 2011; Taboada & Guthrie, 2004). Overall, participants’ strategic competence far exceeded their strategic awareness. Although they rarely named or indicated conscious awareness of the specific strategies that they were employing or their reasons for employing them, students in our sample showed sophistication in their interactions with the texts. Their interactions were indeed strategic in the sense that they were intentional and goal-directed in their pursuits of coherent understandings.

Second, all participants demonstrated in their verbal data that they successfully constructed a propositional
TABLE 8
Case Summaries of Selected Think-Aloud Participants' Reading Processes

<table>
<thead>
<tr>
<th>Reader</th>
<th>Working toward coherent situational understanding</th>
<th>Integration of multiple texts</th>
<th>Epistemic indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Self-explanation</td>
<td>Monitoring and questioning</td>
<td>Intertextual linkages</td>
</tr>
<tr>
<td>A</td>
<td>Minimal use of self-explanation; primarily restates ideas from the text to form a propositional understanding of the broad ideas in the text; restatements somewhat paraphrased but not strongly enough to suggest that the reader is building on background knowledge to achieve coherence</td>
<td>Few indications of mental monitoring</td>
<td>Primarily describes specific topic similarities across texts, with some evidence of comparison of different perspectives represented in the texts; no credibility or authorial comparisons made across the texts other than referring to text 1 as being largely opinion-based compared with text 2</td>
</tr>
<tr>
<td>F</td>
<td>Relies primarily on questioning and monitoring to build a coherent situational understanding of the Pluto issue; other forms of self-explanation not prominent</td>
<td>Expresses tremendous confidence in her understanding of the texts (e.g., “It wasn’t hard to understand what the author was trying to get across”); verbal data contains frequent examples of spontaneous monitoring in the form of questioning, particularly questions expressed as “I wonder…” statements; asks frequent unprompted questions about disciplinary processes described in the text (e.g., “Why didn’t they ask other people’s input?”) and other information in the text (e.g., “Is it possible to send a spacecraft to Pluto’s surface?”); questions indicate monitoring of not only her own understanding but also the accuracy of the text and the viability of the IAU’s decision-making process</td>
<td>Intertextual linkages focused on similarities and differences in content across the texts, whether or not the authors are astronomers, and the perspectives represented in the texts; explicitly acknowledges that by attending to the different opinions expressed in the three texts, she might be able to form her own opinion about Pluto</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued)
understanding of the text content, that is, accurate literal understanding of the ideas explicitly stated in the text. In addition to self-explanation and monitoring (described next), participants were found to prominently use paraphrasing or repeating text information to support and reveal these understandings. With the exception of

<table>
<thead>
<tr>
<th>Reader</th>
<th>Self-explanation</th>
<th>Monitoring and questioning</th>
<th>Integration of multiple texts</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>Self-explanation is prominent in texts 1 and 3; paraphrasing is more prominent in text 2 as the reader talks herself through many of the details described in that text (e.g., “Because of the third rule, which is it has a lot more mass than the other objects that orbit the sun in its same path, Pluto can’t be a planet”); builds a strong situational understanding of the Pluto issue</td>
<td>No evidence of mental monitoring</td>
<td>Intertextual linkages mostly include comments about topic similarity of the texts (e.g., “It also talks about Eris and whether or not it should be a planet”), with some comparison of discursive structure of the texts (e.g., “[Text 1] it’s like in chronological order, and then [text] 2, it’s just starting by saying how it was discovered, then how it went. And the third one is just listing”)</td>
</tr>
<tr>
<td>J</td>
<td>Both self-explanation and questioning are prominent and spontaneous, resulting in a coherent situational understanding of the Pluto issue; attends closely to the specific scientific reasoning for Pluto’s demotion; focuses his self-explanations on the IAU’s decision-making process and the reasons behind Pluto’s ineligibility (e.g., “It seems that the first two conditions of their [IAU] qualifications have been met, but then the third isn’t, because it already mentioned how small Pluto is, and it doesn’t seem that it has enough mass”; “I’m noticing the asteroids may have something to do with what they’ve discovered about Pluto, like the size difference between them and planets” [in text 2])</td>
<td>Engages in a large amount of spontaneous monitoring in the form of statements beginning with “I’d like to know...” and “I’m wondering...”; these statements indicate an emphasis on monitoring his own understanding of the text content and the knowledge creation processes described in the text (e.g., “I’d like to know more about the New Horizons spacecraft and its mission”; “What might be considered the difference between asteroids and other objects?”)</td>
<td>Intertextual linkages focus primarily on comparing the authorial perspectives represented across the texts; also links texts based on topic similarity and makes comparisons across texts related to the depth of information provided on similar topics (e.g., “This text goes more in depth about Eris and how they named it than the other text did”)</td>
</tr>
</tbody>
</table>

Note. IAU = International Astronomical Union.
reader A, participants also showed evidence of constructing a coherent situational understanding of the issue of Pluto’s reclassification, an understanding that integrates propositional representations with prior knowledge and information from other segments of the text (Côté & Goldman, 1999).

Third, even though students’ verbal data indicated that they came to understand the basic information about the IAU’s decision-making process, Pluto’s discovery and naming in the past century, the discovery of new distant objects that complicated the status of Pluto, and the controversy over the Pluto decision, only the participants who had previously scored highest on the expository comprehension measure demonstrated understanding of the specific scientific reasoning (i.e., the neighborhood rule) for Pluto’s demotion. Further, concepts in the texts that were related to disciplinary inquiry and knowledge creation transcending the specific case of Pluto (e.g., self-correction as an important part of scientific inquiry over time, the classificatory nature of the work of astronomy, the prevalence of uncertainty in science) were not commonly expressed in the think-aloud data.

Strategically Working Toward Coherent Situational Understandings Through Self-Explanation and Monitoring

In this section, we describe patterns in the ways that students strategically made meaning with the texts. Think-aloud data reveal that all 10 participants engaged in self-explanation, defined as a sense-making process that “brings new information to the focal segment, including relating it to prior knowledge or information in other segments” (Goldman et al., 2012, p. 363). However, the prominence of self-explanation varied across participants, and with the exception of readers H and J, self-explanation was generally subordinate to paraphrasing or repeating text information in the readers’ interactions with the texts (summarized in Table 8). Nonetheless, we interpret the relative prevalence of self-explanation as an indication that readers were constructing situational understandings of the texts that extended beyond the textbase representations (Kintsch, 2004) made evident in their paraphrasing of content.

When using self-explanatory reasoning, participants connected ideas across different parts of single texts, across separate texts, and to their prior knowledge of the topic to form integrated understandings of the text set. For example, as summarized in Table 8, reader J engaged in self-explanation after reading the definitional features of planets endorsed by the IAU in text 2. Rather than simply restating the three planetary features, he paused and reasoned about the specific factor that resulted in Pluto’s elimination, making explicit references to the information he gained in text 1 and the earlier paragraphs of text 2.

Participants’ situational understandings were also evident through their monitoring of their understanding of the text information, which often occurred in the form of questioning. The ways monitoring was used to support meaning making varied across participants and closely aligned with their performance on the expository comprehension measure. Half of the participants (readers A, C, D, E, and H), mostly those who scored lower on the comprehension measure, either engaged in very little mental monitoring or focused their monitoring almost exclusively on their understanding and pronunciation of individual words (e.g., technical vocabulary, proper nouns). Reader C was the most logocentric, or word-driven (Jiménez, García, & Pearson, 1995), in his monitoring compared with his peers. He was found to engage in very little spontaneous monitoring, and when he was asked by the researcher to identify anything that was challenging about the text, he listed words that he found difficult, including academic terms such as colleagues, underworld (from text 1), controversial, disagree, universe, classification (from text 2), and unpredictable (from text 3). Like his peers who also engaged in logocentric monitoring, he did not articulate any specific plans for resolving his difficulties with these terms.

Higher performing students tended to engage in metacognitive monitoring at a meaning-focused level that might better support coherence building and intertextual integration. For example, reader F, who engaged in more monitoring than her peers, spontaneously stopped twice in the following excerpt to express her “wonderings” about text 1.

This isn’t the first time astronomers have had to rename an object after making a mistake. ‘If astronomers have made a mistake before, I’m wondering if they have made many other mistakes.” In the early 1800s, scientists thought they had discovered four new planets between Mars and Jupiter. It took many years before scientists agreed that these four planets weren’t planets at all. “Hmm, if the planets that they’re talking about is part of our solar system now.” Now we know there are perhaps millions of small objects orbiting the sun between Mars and Jupiter, but now we call them asteroids.

After reading this segment, the student was prompted to explain anything else she was noticing or thinking about the text, and she responded by asking, “When did they start calling these other smaller objects asteroids?” This response suggests that she had located an answer to her previous question and is using this form of questioning to support her ongoing understanding of the concepts in the text related to scientific inquiry and its potential fallibility. Readers I and J were found to use similar forms of meaning-focused monitoring. Both of
these readers spontaneously monitored their understanding of the text content and the disciplinary knowledge creation processes described in the texts. These participants’ verbal data indicate high levels of monitoring of their own thinking and understanding of the text content, but there was little evidence that students were monitoring the credibility of the texts, the credibility of the authors, or the quality of the authors’ arguments.

Integration of Multiple Texts Through Intertextual Linkages and Theorizations About Authorial Identities

The next aspect of multiple-text comprehension that was salient in our analyses was the extent and form of students’ integration of ideas across the three texts. Participants made some efforts to avoid treating each text as a separate island (Bråten & Strømsø, 2006; Wineburg, 1998) by discussing connections or differences that they noticed across texts. However, with some exceptions, verbal explications of intertextual linkages primarily occurred when participants were specifically prompted to describe their process of combining information across texts.

When readers made intertextual linkages, they were most likely to comment on topic similarity (or dissimilarity). All participants identified text information that was repeated across the texts, and most were able to pinpoint specific topics that were mentioned in one text and then elaborated on with more depth in another text. For example, when reading text 2, reader I had this to say about the similarities that she was noticing with text 1:

In the first text, the author was also really upset that Pluto is no longer a planet, and in this one [text 2], it also says that Pluto is no longer a planet. In the other text [1], the author said that scientists always make mistakes and always correct them. In this text [2], it says that science must always be open to correcting its mistakes.

These recognitions of topic similarity were particularly evident when students made comparisons between texts 1 and 2, which contained overlapping and complementary information.

In addition to linking the texts based on content similarity, six of the participants (readers A, B, F, G, I, and J) paid attention to the differences in authorial intent and/or perspectives across the texts. For example, when summing up what he noticed about the three texts, reader A had this to say:

In text 1, you have one person’s opinion and how it [Pluto] was named and how they wanted it to be a planet. And text 2 talks about how the IAU didn’t want it to be a planet. And text 3 combines both and ends up supporting number 1 more than it does number 2.

Finally, intertextual linkages related to text structure and organization tended to focus on external structural differences (e.g., presence/absence of headings within a text), suggesting that participants mostly attended to surface structure when making these comparisons. Two participants (readers B and H) noted differences in discursive structure when they commented on the chronological structure of texts 1 and 2 compared with the nonordinal list structure of text 3. However, we did not find evidence in the verbal data that students were using their structural awareness to support their meaning making while interacting with the texts.

Think-aloud participants’ intertextual integration was also informed by the identities that they constructed for the authors of the texts, which suggests that they were using textual metadata as a supplement to content similarity to form integrated understandings of the Pluto issue. However, they were more likely to acknowledge and construct an identity for the author of text 1 (the personal account of an astronomer’s childhood affection for Pluto) than for the other two texts. Text 1 was the most personalized and provided the most transparency about the author, including a name and information about the author’s profession.

For example, when stopping to verbalize what he was noticing about this text, reader B theorized that “this scientist is very new to astronomy, and [s]he is compassionate about astronomy and wants to know more about astronomy.” He went on to speculate that the author is currently in college and will be an astronomer in the future. Although his theorization is not completely accurate, it suggests that the student was constructing an identity for the author that fit his understanding of someone who is analyzing data to learn more about planets.

Participants generally acknowledged only a vague authorial presence when reading and talking about text 3 (the explicit argument for Pluto’s reinstatement as a planet). Although they often attributed the ideas in the text to “the authors” or “they,” they did not routinely theorize about who the fans of Pluto might be or what credentials or expertise they might hold. For example, when asked to describe the author of text 3 after reading all the texts, reader B simply stated, “The author was against Pluto being a dwarf planet.” Reader F came closest to critiquing and theorizing an identity for the authors of text 3 when she commented, “It’s probably not written by an astronomer because…[they] would have talked more professionally.” She went on to suggest that the fans of Pluto do not appear to have much “authority over this subject.” With the exception of this minor critique, the participants did not routinely evaluate the reasonableness of the argument made by the authors of the texts.
text 3, which was intentionally written to incorporate unsophisticated argument techniques.

In contrast to the author-centered interactions that participants had with text 1 and, to a lesser extent, text 3, most participants framed text 2 (the textbook-like account of the IAU’s decision-making process) as voiceless. Their comments about this text tended to focus on what "the text" was saying while dismissing or ignoring the potentially biased person behind the text. Notably, all participants picked up on the stance/argument represented in texts 1 and 3, even those who did not explicitly theorize an authorial identity for these texts. However, text 2 was largely regarded as representing a neutral stance on the Pluto decision. Reader E even went so far as to conclude that text 2 "is just information that doesn't really have a point of view."

Unlike most of the other participants, reader J theorized an identity for the author of text 2 and commented on the pro-IAU perspective represented in this text. When he initially began reading text 2, he commented about the neutrality of the text in a way that was similar to his peers:

Pluto was discovered in 1930 by an astronomer in Arizona in the southwestern United States. "It seems from a tone of voice in this that this will be mainly an instructional and an educational view of Pluto so far, not as opinionated."

Yet, as he continued reading, he amended his neutral view of the author and pinpointed a specific portion of the text in which the author's stance becomes more transparent:

The IAU decision has been controversial. Many people are upset that Pluto is no longer a planet. Anyone who disagrees with the IAU decision should remember the wise words of a NASA scientist who wrote, "The universe is what it is, not what we want it to be, and science must always be open to correcting its mistakes." "In this part, the author seems to be indicating that he's for their decision to be qualified as a dwarf planet, because he says that anyone who opposes this decision should remember that the scientist has made a point in favor of it."

He went on to conclude that the "second author seems to have studied Pluto for a long time and learned about its specific qualification to the IAU, and he supports their decision."

**Indicators of Epistemic Thinking Related to Scientific Knowledge Creation**

Given the difficulties that we encountered in quantitatively measuring the epistemological dimensions of students’ disciplinary understandings (i.e., how knowledge is built and revised), we looked for indicators in the verbal data that might illuminate their views on the structure and stability of knowledge. These indicators included statements about the readers’ levels of confidence in the truthfulness or objectivity of scientific statements made in the texts.

Participants demonstrated a tendency to defer to institutionalized authority when evaluating credibility of the texts. The information in the text with the most opaque authorial perspective and closest alignment with the IAU’s process (text 2) was considered unquestionably true by nearly all participants. For example, reader F stated, “The second text is most believable because it shows what the IAU felt about planets or Pluto being a planet.” Nearly all participants expressed a similar positive evaluation of text 2 when prompted to comment on the trustworthiness of the texts. The only exception was reader E, who commented that text 1 is likely the most true because it recounts a personal experience. In her own words, “Usually when you’re talking about your experience, it’s a fact. It’s supposed to be factual information.”

Relatedly, participants also had a tendency to make sharp distinctions between facts and opinions when discussing their views on the three texts. This tendency is exemplified in the comments of reader I, who positively evaluated the epistemic status of the claims in text 2 because of the presence of “facts”: “I find the second text more believable because it is based on true facts and there isn't that many opinions, but in the first and third texts, there's a lot of opinions, and [they] list their own thoughts.”

The information in text 1, which aligns nearly perfectly with the information deemed factual in text 2, was deemed less reliable (and more opinionated) because of the personalized narrative style of the text. The possibility that verifiable scientific facts might be interspersed within an opinionated text, and vice versa, was only evident in the think-aloud data from reader J. He acknowledged that despite the ostensible neutrality of the second text, the author implies a personal opinion on the Pluto issue, as indicated in the author's endorsement of a pro-IAU quote from a "wise NASA scientist."

**Summary of Findings**

In response to our first research question—What are the relative contributions of the hypothesized predictors to monolingual and bilingual middle school students’ comprehension of multiple expository science texts?—we found that English-language proficiency and prior content knowledge were independent predictors of expository comprehension in this sample, but strategy expertise and epistemological beliefs were not. English-language proficiency was the strongest predictor, explaining twice as much variance in comprehension as content knowledge. We also found evidence that
the relationship of English-language proficiency to expository comprehension was dependent on participants’ language background, with the relationship being stronger for proficient and emergent bilingual students than for monolingual students.

In response to our second question—What disciplinary meaning-making features are evident in students’ interactions with multiple texts on a scientific topic?—bilingual students were found to engage in processes considered characteristic of expert disciplinary reading, including metacognitive monitoring, theorizing authorial identity, and intertextual integration as they worked toward situational understandings of the issue described in the texts. However, when monitoring, students tended to focus more on their own concept-level understandings than on evaluative monitoring of text and authorial credibility. Intertextual integration focused primarily on topic similarity, with some attention to varying perspectives. Authorial identity theorization was more likely when the author was more transparent. Furthermore, the verbal data provided a window into the ways that students’ views of knowledge played out in their interactions with the texts. Students tended to make a sharp distinction between facts and opinions, with facts assumed to be true information in a text that is not representative of a specific perspective, and opinions assumed to be attributable to an imaginable author.

Discussion and Implications

The present study is one of the few considering the predictors and processes of multiple-text reading with middle school readers and, to our knowledge, is the only study examining these processes using a mixed-methods design and including both monolingual and bilingual readers of this age group. This study affirms the importance of language and content knowledge in reading comprehension of science texts and places traditional strategies in a lesser position relative to these two predictors. Furthermore, the contrast in results between the regression and think-aloud analyses prompts questions about the ways that measurement tools influence our understanding of the role of strategies and epistemic thinking in comprehension processes. While being appropriately cautious about generalizing from our sample to broader communities of students, we discuss the implications of our findings in the sections that follow.

The Limited Role of Strategy Knowledge Relative to Other Predictors

Expertise in the use of cognitive and metacognitive strategies has become a central learning goal in reading comprehension instruction, but our findings suggest that students’ knowledge of discipline-generic strategies used for single-text reading is not a strong contributor of multiple-text comprehension, particularly for students who are still developing in their English-language proficiency. Although we did not test a specific instructional model in this correlational study, the results suggest that if students in this sample were to make gains in their knowledge and awareness of traditional strategies, there would be minimal change in their successful construction of understanding when reading multiple expository texts.

Other studies have found similarly small relative contributions of strategy knowledge in explaining variation in expository comprehension. In a sample of linguistically homogeneous high school students, Samuelstuen and Bråten (2005) found that although self-reported use of organization and monitoring strategies was significantly related to expository comprehension, topic knowledge was a stronger predictor. O’Reilly and McNamara (2007) found that for high school students across four socioeconomically diverse contexts, strategy knowledge as measured by the MSI was less strongly related to science comprehension compared with science knowledge and general reading skill.

The present study corroborates these previous findings and also prompts new questions about the importance of strategy awareness that should be pursued in future work. Students’ general awareness of strategies (as measured by the MARSi) was not found to relate to comprehension in the bivariate analysis or the multiple regression model. Further, in the think-aloud data, students’ sophistication in strategic meaning making far exceeded their conscious strategic awareness. They demonstrated emergent expertise in using important strategies for monitoring and supporting their construction of meaning, but they did not regularly put a label on their strategy use or explicate how they were using strategies. The fact that across two data sources, strategy awareness did not emerge as an important feature of comprehension should prompt future research on the dimensions of strategic expertise that make a difference in readers’ constructions of meaning from text.

A related question that emerges in our analysis is whether it is appropriate to measure strategy expertise using survey instruments. Some previous research employed online measures of strategy expertise—that is, measures in which students demonstrate strategic expertise at specific points while reading a text—and found stronger relationships with comprehension (e.g., Cain et al., 2004). Cognitive and metacognitive strategies are fleeting, in-the-moment acts that support monitoring and meaning in the act of reading (Afflerbach, Pearson, & Paris, 2008), and they are
difficult to capture outside of actual reading events. It is possible that strategy awareness, as measured by a self-report survey, was not found to relate to comprehension in the present study because this survey measured familiarity with traditional single-text strategies rather than a genuine tendency to spontaneously use these strategies. Strategy knowledge may have been only weakly (and nonsignificantly) related to our comprehension measure in the regression model for the same reason. A major contribution of the present study is the use of think-aloud methodology to show that students demonstrated competence in a repertoire of intertextual meaning-making strategies that were not assessed on the survey measures.

**Conditional Relationship of English-Language Proficiency and Reading Comprehension**

English-language proficiency was the strongest predictor of expository text comprehension examined in the present study. This finding corroborates other recent studies that highlighted the strong role of vocabulary in comprehension (Language and Reading Research Consortium, 2015; Liebfreund, 2015). Our finding that English-language proficiency was more strongly associated with expository reading comprehension among bilingual students than monolingual students corroborates previous research comparing the contribution of language proficiency across monolingual and bilingual samples (Drop & Verhoeven, 2003; Geva & Zadeh, 2006). The strengthened relationship observed in the present study between comprehension and language proficiency for bilingual learners could be indicative of a linguistic threshold (Cummins, 1979) that learners need to reach in English before the facilitative effects of strategy knowledge and content knowledge become evident. However, this possibility is complicated by the fact that bilingual status was only weakly related to English-language proficiency in this sample ($r = -0.175$). In fact, proficient bilingual and monolingual readers scored similarly on our measures of vocabulary, syntax, and morphology. Also, emergent and proficient bilingual readers were found to have similarly strong associations between language and comprehension. These patterns suggest that this difference in relationship might be due to other factors besides language proficiency.

For instance, bilingual readers with high L2 proficiency may be more heavily leverage their linguistic knowledge as a resource in comprehension compared with monolingual readers. This aligns with the qualitative findings of Jiménez et al. (1995), who found that a proficient Spanish–English bilingual reader showed different reading processes compared with a monolingual English reader and a less proficient bilingual reader. The proficient bilingual reader demonstrated awareness of the similarities and differences between Spanish and English and exploited the relationship between the two languages to help her comprehend English text. Unpacking the possible mechanisms that may account for the complex relationship between language proficiency and L2 reading comprehension is an important direction for future work. For example, follow-up studies that directly compare the think-aloud processes of monolingual, proficient bilingual, and emergent bilingual readers as they interact with disciplinary texts could provide valuable insights.

**Epistemology and Reading Comprehension**

Unlike previous work that suggested the existence of a relationship between epistemological beliefs and comprehension (Bråten & Strømsø, 2006; Schommer, Crouse, & Rhodes, 1992; Schommer-Aikins & Easter, 2009), we did not find such a relationship in this sample. The lack of a relationship in the present study may be due to the developmental nature of epistemological sophistication in young adolescents (Schommer-Aikins et al., 2000), which makes it challenging to identify differences in epistemic thinking across readers. Indeed, the think-aloud cases in this study demonstrated minimal variation in their tendency to privilege voiceless, nonpersonalized knowledge. They generally demonstrated beliefs about knowledge that previous researchers have referred to as naive (Bråten & Strømsø, 2006) or absolutist (Barzilai & Zohar, 2012) epistemologies.

Alternatively, it is possible that we did not find a strong relationship between epistemology and comprehension in this study because epistemic beliefs indirectly affect expository text comprehension in ways that are not likely to show up in the model that we tested. Evidence for this possibility comes from Bråten, Anmarkrud, Brandmo, and Strømsø (2014), who found that secondary students’ beliefs about the need to corroborate knowledge claims in science were not directly related to multiple-text comprehension, but they were related to process variables (effort, strategies, and interest), which were related to comprehension. Previous research has also shown that epistemic beliefs may affect some features of comprehension processes more than others (e.g., Barzilai & Zohar, 2012). These more complex relationships were not examined in the present study and should be explored in future research.

It may also be the case that even if there is a relationship between epistemological beliefs and comprehension, other factors may override this relationship for linguistically diverse middle school students. In our sample, one epistemology scale (structure of knowledge)
was modestly correlated with comprehension, but this relationship was reduced in the regression model once other factors were taken into account. Additional work is needed to further examine this relationship, particularly among young adolescent readers who are still developing in their L2 proficiency.

Theorizing About Authors

Finally, our study supports a need for more research on how young adolescent readers from linguistically diverse communities theorize about authors when reading informational texts in science. As posited in theoretical accounts of successful multiple-text reading (Bråten et al., 2011), readers must engage in an interaction with both the information in the texts and the author(s) behind the production and dissemination of the texts. Our findings from the think-aloud analysis suggest that the young adolescent readers in this sample may be just beginning to interrogate authors in ways that would support critical reading of science texts consistent with this perspective.

The fragilities that we observed in how readers interacted with metadata related to authors could suggest a need for middle school readers to have more instructional opportunities to interrogate authors and take a critical stance as part of their meaning-making processes. Although we did not collect any data in the students’ classrooms, the literature suggests that the epistemic climate in many U.S. classrooms is often monologically organized, with knowledge transmitted by a known authority to learners who are expected to passively absorb it (Aukerman & Schuldt, 2015). This is in contrast to more dialogically organized classrooms, in which knowledge is framed as ambiguous, collaborative, and subject to revision (McElhone, 2014). Our findings point to a need for future research exploring the ways in which the classroom epistemic climate affects variation and sophistication in the way readers interrogate the authority of authors in their text-based knowledge construction processes.

Furthermore, our findings suggest that the texts themselves may influence the degree to which readers interact critically with authors. When the author’s identity and position were made more explicit (with a byline in the case of text 1 and a group affiliation in the case of text 3), participants were more likely to attribute ideas to a specific perspective and to feel emboldened to question this perspective. When the author’s identity and stance remained opaque, students were more likely to accept the information as neutral and objectively true. This pattern of differences in critical interrogation of information presented in science-related texts warrants examination in future work.

Limitations

This study has a few limitations. First, the epistemology measures had lower reliability than desired, and we are careful to take this into account when interpreting the findings. Also, we used only two texts in the quantitative measure of multiple-text comprehension and only three texts in the think-aloud protocol. This makes the version of comprehension examined in this study less complex than the authentic multiple-source comprehension processes that middle school students may experience outside of school. Our choice to include only bilingual students in the think-aloud sample limits our ability to make comparisons of the multiple-text reading processes between bilingual and monolingual students, which could be an important direction for future work. Additionally, we administered generic measures of strategy knowledge and awareness that may not capture the full range of strategic interactions used in science reading or multiple-text integration. Strategy knowledge and awareness measures specifically designed for multiple-text comprehension might have had stronger relationships to the outcome measure. Finally, our modest sample size precluded the examination of more complex multivariate and interactive relationships that should be studied in future work on this topic.

Conclusion

This study shows that students’ comprehension of science texts was related to their English-language proficiency and prior content knowledge, but not their strategy expertise and epistemological beliefs; these relationships differed for monolingual and bilingual students. Although traditional strategy expertise was not a unique predictor of comprehension, students showed evidence of strategic meaning making in their interactions with multiple texts, suggesting that the complex processes of multiple-text comprehension are easily within striking distance of young emergent and proficient bilingual adolescents. This trend suggests to us that students like those in our sample would be able to handle increased opportunities for complex disciplinary reading, especially if appropriate scaffolds are in place to help them leverage their linguistic knowledge and emerging content knowledge in science. Additional work with readers of this age is needed to further examine the relationships made evident in our data and to examine the instructional factors that support emergent and proficient bilingual students in expository reading comprehension.
NOTES
This research was supported by a faculty research award from the College of Education and Human Development at The University of Texas at San Antonio.

1 We use the terms EL and emergent bilingual interchangeably in this article, although we prefer the latter term because it represents a more asset-oriented view of where these learners are headed in their language development. Throughout this article, we use the term proficient bilingual to refer to bilingual students who are not considered by their schools to be in need of English-language support and emergent bilingual for those who are identified as needing this support.

2 In the original version of Schommer-Aikins et al.'s (2000) measure, this item was included in the stability-of-knowledge subscale, but in the present study, it correlated more highly with the structure-of-knowledge items.

3 The negative direction of this relationship contradicted our expectations. The scale for this measure was structured so higher scores indicate increased alignment with a view of knowledge as dynamic and evolving. However, the two items (listed in the Methods section) on this scale both refer to the work of scientists. It is possible that a low score on this measure (high agreement with the two items) reflects an endorsement of the power of science and scientists to reliably work toward accurate understandings. Ascribing to this endorsement of science could conceivably be associated with improved understanding of the Pluto texts. Although our texts were designed to highlight the uncertainty of scientific knowledge, they could be read as privileging a view of scientific knowledge moving with increasing accuracy toward a final endpoint (i.e., the best way to classify the objects in the solar system). This epistemological position would more closely align with lower scores on the stability-of-knowledge measure (i.e., higher agreement that scientists can eventually reach the truth) than with higher scores.

4 In this example and similar examples, bolded text is the segment of text that the student is reading, and nonbolded text in quotation marks is the student’s talk.

REFERENCES


Brätten, I., Annmarkrud, Ø., Brandmo, C., & Stroømo, H.I. (2014). Developing and testing a model of direct and indirect relationships between individual differences, processing, and multiple-text comprehension. Learning and Instruction, 30, 9–24. doi:10.1016/j.learninstruc.2013.11.002


comprehension. *Journal of Educational Psychology*, 97(2), 246–256. doi:10.1037/0022-0663.97.2.246


Submitted March 10, 2015
Final revision received July 11, 2016
Accepted July 18, 2016

DENNIS S. DAVIS (corresponding author) is an associate professor in the Department of Teacher Education and Learning Sciences at North Carolina State University, Raleigh, USA; e-mail ddavis6@ncsu.edu. His research focuses on reading comprehension instruction and assessment in culturally and linguistically diverse classrooms. He was a faculty member at The University of Texas at San Antonio when this study was conducted.

BECKY HUANG is an assistant professor in the Department of Bicultural and Bilingual Studies at The University of Texas at San Antonio, USA; e-mail becky.huang@utsa.edu. She specializes in L2 acquisition and language assessment.

TANISHA YI is a recent graduate from the master’s program in the Department of Interdisciplinary Learning and Teaching at the University of Texas at San Antonio, USA; e-mail tanishayi.ty@gmail.com. Her research focuses on critical instruction throughout informal education.

**Supporting Information**

Additional supporting information may be found in the online version of this article on the publisher’s website:

- **Appendix: Texts and Items: Measure of Expository Comprehension With Multiple Texts and Think-Aloud Protocol**