Running head: DON'T KNOW

# What You Don't Know Won't Hurt You, Unless You Don't Know You're Wrong

Tenaha O'Reilly[1], John Sabatini[2], Zuowei Wang[1]

[1]Educational Testing Service, Princeton, New Jesery

[2]The University of Memphis, Memphis, TN

Reading Psychology, Volume 40, 2019 - Issue 7

Published Online: 27 Sep 2019

Peer review policy:

https://www.tandfonline.com/action/journalInformation?show=aimsScope&journalCode=urpy20

## Acknowledgments

This research was supported by the U.S. Department of Education Institute of Education Sciences (Award Nos. R305A150176 & R305F100005) to Educational Testing Service. The opinions expressed here are those of the authors and do not represent the views of the U.S. Department of Education or Educational Testing Service.

## Abstract

The purpose of this investigation was to determine the impact of students’ background knowledge and how they utilized “don’t know” affordances to comprehend and learn from text. In two studies, over 8,000 middle and high school students interacted with a content-area learning environment in which they answered a series of background knowledge questions before they completed a unit on the same topic. Students were given the opportunity to indicate they “did not know” the answers to the knowledge questions. Higher knowledge was related to higher understanding, and the use of the “don’t know” option further explained variability in students’ understanding of the sources beyond background knowledge. When responding to knowledge questions, students who selected incorrect options before the task understood less and were less likely to learn content when given the opportunity compared to students who indicated they did not know. Thus, low knowledge students were still able to comprehend and learn as long as they acknowledged they lacked background knowledge. One’s comprehension and learning can be facilitated or impaired, depending upon the veracity of their knowledge, and whether students choose to acknowledge their lack of background knowledge. Implications of this work are discussed in terms of learning and instruction.


Keywords: content-area learning environment, topical knowledge, don’t know option, learning, understanding

# What You Don't Know Won't Hurt You, Unless You Don't Know You're Wrong

## 1 Introduction

In the age of the internet, smart phones and fake news, what it means to understand and learn in 21st century literacy environments has changed over the past 20 years (Alexander, 2012; Goldman, Britt, et al. 2016; LaRusso et al. 2016; Leu, Kinzer, Coiro, Castek & Henry, 2013; NGA & CCSSO 2010). Technology and the availability of the internet have expanded the opportunity for accessing information while simultaneously placing demands on attention and cognitive resources for processing (Magliano, McCrudden, Rouet, & Sabatini, 2018). Moreover, the quality of information is often suspect (Metzger, 2007), requiring people to identify information that is incorrect, or just pure propaganda. On balance, reading and learning in modern environments require a diverse set of integrated skills. For instance, students often need to engage in complex reasoning and perspective taking (LaRusso et al., 2016) as they evaluate and integrate multiple (Organisation for Economic Co-operation and Development, 2016) and digital sources (Leu et al., 2013) to achieve their reading goals (Britt, Rouet, & Durik, 2017). This shift has also been reflected in the U.S. Common Core State Standards (NGA & CCSSO 2010), where reading comprehension has expanded to include content area and disciplinary reading literacy (Goldman, Britt, et al., 2016).

However, as decades of research have indicated, the ability to understand content is dramatically impacted by what one already knows about the topic (Shapiro, 2004). In some cases background knowledge can facilitate comprehension when it is correct and relevant to the topic of the sources (Ozuru, Dempsey, & McNamara, 2009). In other cases, background knowledge can be incorrect and impede understanding (Kendeou & van den Broek, 2005). To compound matters, incorrect knowledge and misconceptions can be difficult to fully overcome

(Meir, Perry, Stal, Maruca, & Klopfer, 2005; Özgür, 2013), even when students are given information contrary to their prior misconceptions (Broughton, Sinatra, & Reynolds, 2010; Diakidoy, Mouskounti, Fella, & Ioannides, 2016; Lassonde, Kendeou, & O'Brien, 2016; van Loon, Dunlosky, van Gog, van Merriënboer, & de Bruin, 2015). Students might have accurate or inaccurate knowledge, but they also might have little or no knowledge about a topic. In cases where a student has little knowledge, they should be aware of their lack of knowledge (Tobias & Everson, 2009) as a first step in the process of "vaccinating them from the prevalent germ of misinformation".

The goal of the current research is to investigate how students' background knowledge impacts comprehension and learning in a digital content learning environment. In particular, we explore how individual differences in acknowledging one's low knowledge on a topic is associated with comprehension (Study 1) and learning as defined by pre and post-test changes in students understanding (Study 2). To create an environment for students to freely acknowledge their lack of knowledge, we provided students with an "I don't know" (IDK) option on each question in a background knowledge test. We hypothesized that students who failed to acknowledge their low knowledge were also less likely to understand (Studies 1 and 2) and learn from content texts (Study 2) than students who acknowledged their low knowledge state (selected the IDK option).

Before describing the current study, we briefly review the literature on the role of background knowledge, and how "don't know" options have been used in the study of reading comprehension.

## 1.1 Background Knowledge and its Effect on Understanding and Learning

Over the past several decades, there has been a growing body of research documenting empirically the influence of background knowledge on students’ understanding (Alexander, Kulikowich, & Schulze, 1994; Alexander, Sperl, Buehl, Fives, & Chiu, 2004; Cromley & Azevedo, 2007; Dochy, Segers, & Buehl, 1999; Fincher-Kiefer, Post, Greene, & Voss, 1988; Hambrick & Engle, 2002; McNamara, 1997, 2001; McNamara & Kintsch, 1996; Murphy & Alexander, 2002; Ozuru, Best, Bell, Witherspoon, & McNamara, 2007; Recht & Leslie, 1988; Schneider, K€orkel, & Weinert, 1989; Shapiro, 2004; Thompson & Zamboanga, 2004; Voss & Silfies, 1996). This literature has shown that background knowledge can improve reading comprehension when the domain knowledge is relevant to the texts (Ozuru et al., 2007) or, in some cases, harm reading comprehension when the background knowledge contains a misconception, conflicts with the text, or is interjected inappropriately during a retell (Kucer, 2011; Marsh, Meade, & Roediger, 2003). For instance, Kendeou and van den Broek (2005) found that students who had prior misconceptions were more likely to recall less information after reading, generated more invalid inferences and fewer valid inferences than students with no prior misconceptions. Thus, background knowledge can facilitate or harm students’ comprehension depending upon the accuracy of the knowledge.

### *1.1.1 How Knowledge Impacts Comprehension: Theoretical Perspectives*

Background knowledge plays an essential role in contemporary theories of comprehension (Kintsch, 1998, 2012), expertise (Alexander, 2003) and learning (Kendeou & O’Brien, 2014). For instance, in Kintsch’s (1998, 2012) Construction- Integration model, background knowledge is central to the process of forming a situation model. The Construction-Integration model

provides an explanation of how readers construct multiple levels of text representation. These representations tend to vary in terms of their level of abstraction in relation to the original text and the knowledge demands required to construct a coherent representation. The representation that is most similar to the text is called the surface level, and it is essentially a temporary verbatim representation of the literal words, phrases, and structures of the text. A more abstract level of representation called the textbase contains the set of propositional structures that preserves the semantic and syntactic relationships or “gist” of the text meaning. One might think that the textbase would be sufficient to accurately capture all the nuances and meaning implied by the text. However, the texts are often not completely specified. Authors assume their readers have basic knowledge on the topic of the text (Beck, McKeown, & Gromoll, 1989). And a reader’s purpose for reading a text may extend beyond the original author’s intent when writing it. As a result, readers need to infer unstated or implied relationships in the text, in order to construct a coherent representation (McNamara, 1997; 2001; McNamara & Kintsch, 1996). In particular, some of these inferences, called knowledge-based inferences require that students integrate their background knowledge with the text to build a deeper, situation model of the text’s meaning (Kintsch, 1998). For instance, in the sentences “The plate fell. There were pieces everywhere”, a reader would have to activate their background knowledge about plates, in order to draw the inference that the “plate was broken”. In particular, this relevant knowledge would entail that the reader knows that plates are fragile and may break when dropped on a hard surface such as a floor. Note the sentences do not contain any explicit information that the plate fell on a hard surface such as a floor, and that it is broken. These two concepts must be inferred through the integration of the reader’s background knowledge and the text.

Background knowledge can also make the process of reading comprehension more efficient. If a reader already knows some of the material being read, the reader can skip this familiar material and focus only on the concepts that are unknown in the text (Tobias & Everson, 1996, 2009). This not only saves the reader time, but also reduces the load on processing and memory. Of course this mechanism requires that students are aware of what they know and what they don’t know (Tobias & Everson, 2009), a key element of the current study.

### 1.2 Acknowledging one’s low knowledge: The use of a Don’t Know option to facilitate subsequent understanding and learning

In the current study, we explored the use of an “I don’t know” (IDK) option on students’ ability to learn and comprehend content from multiple text sources. This decision was motivated by a growing body of research that suggests the reliability of a measure can be improved by providing students with an “unsure” or “don’t know” option (Courtenay & Weidemann, 1985; Muijtjens, van Mameren, Hoogenboom, Evers, & van der Vleuten, 1999; Pennington, Pachana, & Coyle, 2001; Ravesloot et al., 2015; Sanderson, 1973; Scoboria & Fisico, 2013; Wakabayashi & Guskin, 2010). For instance, Scoboria and Fisico (2013) found that forcing witnesses to provide answers to every question lead to inaccurate and fabricated accounts of events. In contrast, providing the “don’t know” to witnesses resulted in less misinformation from them. In other words, allowing for an acknowledgement of no knowledge (don’t know) is more productive than encouraging guessing behavior and possibly the adoption of incorrect knowledge when the participant is uncertain of their low knowledge state. Similarly, Wakabayashi and Guskin (2010) argued that random guessing in a forced-choice task only adds measurement error to the score. Therefore participants should be allowed to accurately report that they are

unsure of certain answers. In a pretest and posttest training design, the experimenters found that adding an “unsure” response improved the sensitivity of the tests. Importantly, students who were more likely to answer “unsure” at pretest were also more likely to answer the posttest questions correctly than students who were confident in their guesses at pretest. The experimenters concluded that providing participants with an option that allows for uncertainty (unsure) helps create “teachable moments”. These studies suggest that providing students with an affordance to indicate their uncertainty to an initial response has the potential for improving the quality of the information gathered from them. Furthermore, students who acknowledge their lack of knowledge may show better learning in subsequent learning activities. In the current studies, we were interested in determining whether providing an IDK response for background knowledge items could have an impact on how well students understand content (Study 1) and how well they eventually learn the content after reading sources that provide answers to the former knowledge items (Study 2). In addition, providing students with the IDK option may be useful in distinguishing those who acknowledge that they have low knowledge (selecting IDK), from those who select wrong answers without acknowledging low knowledge. Importantly, we add to the literature by extending the IDK technique to a complex, content learning environment that resembles real-life learning and problem solving activities in the digital age. The learning environment not only measured students’ background knowledge, but also supported and tracked students’ understanding as they engaged in deeper purpose-driven, multiple-source comprehension tasks, as described next.

### 1.3 Building content knowledge from text sources

Students in the studies reported here learned content from text sources in a web-administered, scenario-based assessment (SBA) session. The SBAs used in this study were inspired by a

“building and sharing knowledge” learning model, which is derived from research in reading and the learning sciences (O’Reilly, Deane, & Sabatini, 2015; O’Reilly, & Sabatini, 2013; Sabatini, & O’Reilly, 2013; Sabatini, O’Reilly, & Deane, 2013). The learning model describes a strategic process that involves the integration of several key components or phases. Before reading, students are encouraged to activate their relevant background knowledge (Bråten, Johansen, & Strømsø, 2017), set learning goals (Alexander, 2012; van den Broek, Lorch, Linderholm, & Gustafson, 2001; van den Broek, Risden, & Husebye-Hartman, 1995), identify relevant information (McCrudden, Magliano, & Schraw, 2011), and ask guiding questions (King, 1995; Ogle, 1986) that set the context for learning. During reading, students are encouraged to understand the text by using a set of strategies derived from effective instructional practices (Goldman, Snow, & Vaughn, 2016; King, 1995; McNamara, 2007; Meyer & Ray, 2011) to construct a coherent mental model of the text content that is consistent with their background knowledge. As they engage with content, students clarify meanings of unknown words and concepts and engage in metacognitive and self-regulated learning (Hacker, Dunlosky, & Graesser, 2009; Zimmerman, & Schunk, 2001). After reading, students consolidate what they read by using a variety of further reading strategies that strengthen the representation in longterm memory (Franzke, Kintsch, Caccamise, Johnson, & Dooley, 2005; McDaniel, Anderson, Derbish, & Morrisette, 2007; Meyer & Wijekumar, 2007; Ogle, 1986). These may include synthesizing, comparing and contrasting, applying knowledge to new contexts. Finally, students convey what they have read in writing, speaking, or other representational formats to reflect communication goals and the intended audience. Collectively, the building and sharing knowledge approach is intended to both model skilled performance and help identify component

skill weakness in students (O’Reilly, Deane, & Sabatini, 2015). Many of the ideas in the model are consistent with modern and effective large scale reading comprehension interventions, which share three key features: they provide meaningful purposes and contexts for reading a diverse collection of sources, they take advantage of social interactions by supporting classroom discussion and team-based learning, and they support content area learning by building and integrating students’ background knowledge and vocabulary (Goldman, Snow, et al., 2016). Overall, the model is instantiated in a goal-directed, web administered, digital SBA session (O’Reilly, & Sabatini, 2013). First, students are presented with a purpose for reading (van den Broek et al., 2001) a collection of thematically related sources that vary in terms of genre (e.g., e-mail, blog, website, policy document), credibility (Metzger, 2007), and relevance (McCrudden, Magliano, & Schraw, 2010; Rouet & Britt, 2011). Then, relevant background knowledge is activated through a series of task activities related to the thematic topic. Next, texts and tasks are sequenced to support the building of background knowledge that is necessary to forming situation models. Typically, the session commences with an introductory text on the thematic topic, along with tasks that support the building of text-based mental models (e.g., summarizing or paraphrasing the text) (Radmacher & Latosi-Sawin, 1995; Youjia, Woods-Groves, Ford, & Nobles, 2014). As the session proceeds, more complex reasoning, perspective taking, and multiple source evaluation tasks are presented to the student that build on knowledge acquired from reading prior texts (Britt et al., 2017; LaRusso et al., 2016). In a nutshell, the sequence of activities in the session provides students with an environment that models effective content learning in a digital learning environment. Thus, students have the opportunity to build their knowledge about a particular topic, following a systematic set of

task activities (see Appendix) built around a scenario context familiar to students. For example, the instructions given to participating secondary biology students in study one were as follows: You are preparing for an upcoming science test. In order to prepare for the test efficiently, you and some classmates decide to form a study group. You are responsible for helping each other identify key concepts, review and learn from readings, and understand scientific terms and data. These directions are followed by introducing the student to simulated, online peers and teacher, and then successive texts and activities that simulate a study group preparing for a test on the topic of ecosystems, invasive species, and related science policy. The scenarios vary in their focus, but each foregrounds the goals of building one’s understanding and knowledge of the content, as well as sharing that knowledge in some re-representation or application of the acquired content knowledge. Thus, the SBA session represents a goal directed opportunity to understand and learn about a topic area. What makes this an assessment is that the development of the students’ knowledge and understanding is successively tracked over the course of the session, creating an evidence trail that can be used as a summative score of their content reading ability (O’Reilly, Weeks, Sabatini, Halderman, & Steinberg, 2014; Sabatini, O’Reilly, Halderman, & Bruce, 2014), or formatively to provide further learning support. Thus, upon completion of the session, student understanding has been evaluated by their responses throughout the tasks. But the student’s focus has been on understanding topical content for a purpose, not merely to answer questions posed in decontextualized, passive voice as is typical of traditional comprehension tests. That is, the scenario is a learning environment for understanding and learning text content, as well as an assessment of student literacy skills. However, understanding and learning about a topic interacts

with one’s background knowledge, and one’s awareness of that knowledge. Which brings us to the question driving these studies.

### 1.4 Overview of the Studies

In the current studies, students were asked a series of background knowledge questions to assess how much they knew about a content area before engaging in a scenario-based assessment (SBA) for understanding and learning about the particular topic. In Study 1, we examined how students’ background knowledge and their utilization of the IDK option impacted their understanding of content area texts in two SBAs. We define understanding here as students’ performance on the SBA tasks while the texts are available. In Study 2, we first replicate findings of Study 1 using a third SBA topic with a different age group, and then we investigated how students’ use of the IDK predicted their learning from the sources. We define learning here as students’ change in performance from pretest (on the background knowledge items) to posttest performance when the same questions are asked again after sources are presented that provide answers to the knowledge questions. Consistent with prior research, we predicted a main effect for knowledge such that students with more background knowledge (Shapiro, 2004) would comprehend more from texts. In addition, we also predicted an effect for the IDK usage over and above the separate contribution of background knowledge, based on prior findings that providing an IDK option can affect how well new information is learned (Wakabayashi & Guskin, 2010). In other words, students who acknowledge their low knowledge state will be more likely to learn the information when given the opportunity. In contrast, students who select wrong options in the knowledge task (incorrect knowledge) will be less likely to learn new information when given the opportunity to correct erroneous knowledge.

### 1.5 Hypothesis

H1: students who correctly answer more background knowledge questions (high knowledge) before engaging in a learning environment, will show better understanding of the content.

H2: between two groups of students who correctly answer the same number of background knowledge questions, those who select more IDK responses (acknowledging low knowledge) will show better understanding of content in the subsequent learning environment than those who select wrong answers.

H3: students who select more IDK responses on the knowledge test, will have more learning gain than students who selected wrong answers to the knowledge questions.

## 2 Study 1

### 2.1 Method

#### *2.1.1 Participants*

A sample of 7,396 grade 9-12 students (mean grade level 10.1, SD .9) who were recruited from 32 schools in two states in the U.S. completed tasks of Study 1. A total of 4,105 students completed the SBA on the topic of Immigration, 3,578 completed the SBA on the topic of ecosystems. All data was collected according to the research institution’s internal review board policy. Due to an agreement with the schools, demographic information was not available at the individual level. However, because this project was part of a larger intervention program (Fancsali et al., 2015), we were able to obtain demographic information for the whole recruitment pool from which the two groups of our participants were drawn. For the whole

recruitment pool of 14,747 students, 49% were female; 14% were English language learners; 56% were eligible for free or reduced-price lunch and 61% were nonwhite students.

### *2.1.2 Instruments*

The current study used two SBAs, one focused on the topic of Ecology (34 items) and the other on United States Immigration in the 19th Century (31 items). The basic psychometric properties and range of scores of these two SBAs were shown to be adequate in other settings (O’Reilly et al., 2014), and the current study with alpha reliabilities of .88 and .87 for the Ecology and Immigration SBAs respectively (see Table 1). The difficulty level was appropriate and no ceiling or floor effect was found. The SBA units included in this study covered a range of item types that targeted a variety of skills, some of which are not typically addressed in traditional reading tasks. For example, application-style items required students to apply what they learned from the reading passages to new situations, while classification items asked students to decide whether statements related to the passages were key concepts, minor details, incorrect, or someone’s opinion. See the Appendix for more information on the range of skills and design of the SBA units. The length of the two primary content passages in the SBA Ecology unit were 814 and 304 words, with Flesch Kincaid (Kincaid, Fishburne, Rogers, & Chissom, 1975) grade levels of 9.8 and 15.4 respectively. The length of the two primary content passages in the SBA Immigration unit were 691 and 410 words, with Flesch Kincaid grade levels of 7.8 and 6.2 respectively. The respective text complexity grade level estimates using the TextEvaluatorVR system (Sheehan, 2016) were 10 and 12 respectively for Ecology, and 7 and 8 respectively for Immigration. The SBA approach has been evaluated in elementary (Sabatini, Halderman, O’Reilly, & Weeks, 2016), middle (Sabatini et al., 2014) and high school populations (O’Reilly et al., 2014). The

results of these studies indicate that the SBAs are reliable (Cronbach's alpha .80 or higher) and students display a range of student understanding with no apparent floor or celling effects. The SBA correlates with the prior year’s high school English Language Arts state test scores ranging from .52 to .68 (O’Reilly et al., 2014). The SBA also correlates with measures of academic vocabulary, complex reasoning, and perspective taking (LaRusso et al., 2016). To date SBAs cover 19 topic areas.

**TABLE 1** Descriptive Statistics of Students’ Performance in SBA Ecology and Immigration

| | Ecology | | | Immigration | | |
|---|---|---|---|---|---|---|
| Measure | Mean | Standard Deviation | Cronbach’s alpha | Mean | Standard Deviation | Cronbach’s alpha |
| Background Knowledge | .66 | .18 | .91 | .64 | .15 | .86 |
| Number of IDK responses | 11 | 9 | NA | 9 | 8 | NA |
| SBA Performance | .44 | .21 | .88 | .57 | .19 | .87 |

*2.1.2.1 Background Knowledge*

Two types of tasks comprised the knowledge test. In the topical vocabulary task of the BK test, students were shown a randomized list of words, about half of which were related to the targeted topic of the SBA form, while the others were unrelated to the topic. Students were asked to choose whether each keyword was related to the topic, unrelated, or could choose the option “I don’t know”. Separate, 44-word lists were constructed for the topic of ecology (e.g., species, habitat) and immigration (e.g., naturalization, visa). For ecology, 26 words were topically related, and of those, 9 appeared subsequently in the Ecology comprehension form. For immigration, 26 words were topically related, and of those, 11 appeared subsequently in the immigration comprehension form. The content-focused items of the BK consisted of multiple

choice items that targeted facts, concepts, principles, or causal mechanisms related to the corresponding topic, that is, Ecology (13 items) or Immigration (8 items). For example, one such item in Ecology asked students to select from four options the correct definition of invasive species. Both topical vocabulary and content knowledge items were presented to the students at the start of the SBA before they read the sources. Students were allowed to give IDK responses for all BK items, and were specifically instructed that their answers to these questions would not count towards their scores on the comprehension test that followed. Alpha reliabilities for the total BK tests were .91 (57 items) and .86 (52 items) for Ecology and Immigration respectively. In addition to calculating the number of background knowledge questions correctly answered, for each student we also calculated the number of IDK responses.

#### *2.1.3 Procedure*

All measures were web-administered to students during a regular, 50-minute class period. The knowledge (including don’t know affordance) section and the content-area learning environment were all part of a single, seamless session. That is, the background knowledge items (with the option to respond IDK) were introduced as the first section of the SBA, though, as noted, students were instructed that their answers would not count towards their final score. This freedom to answer, or not to answer, the knowledge questions was under the student’s control and represented the heart of the IDK affordance used in this study.

### 2.2 Results

Table 1 shows students’ performance on the background knowledge section and content learning section of the two SBA units. Students on average had some knowledge on both topics, correctly answering 66% and 64% of the background knowledge questions on the two SBA units respectively. Students also showed sizable variability on SBA performance in both units,

with the Immigration unit relatively easier than the Ecosystem unit. On about 20% of the background knowledge items students selected the IDK option (11 out of 57 for Ecosystems and 9 out of 52 for Immigration). This shows that students were responsive to the IDK option.

The correlation between knowledge and performance in SBA showed relatively the same magnitude across the two SBA forms, r.48 and .49 for Ecology and Immigration respectively. This supports the first hypothesis and replicates previous studies demonstrating students with higher background knowledge understand the learning material better. To test the second hypothesis that among students who have comparable amount of correct knowledge, those who select IDK responses will show better understanding, we performed hierarchical regression using students’ background knowledge and number of IDK responses to predict SBA performance. In the first regression step, students’ background was entered into the regression model; then in the second regression step the number of IDK responses was entered. Table 2 shows that data obtained from the two SBA units produced comparable regression results. In the first regression step, students’ background knowledge positively predicted 23% to 24% of the variance in students understanding. In the second regression step, after the effect of background knowledge was controlled, the number of IDK responses was positively related to understanding, explaining an additional 6% of variance beyond the effect of background knowledge. Thus, the second hypothesis was supported.

**TABLE 2** Hierarchical Regression Predicting SBA Performance by Background Knowledge and IDK Selections

| | | SBA Ecosystems | | | SBA Immigration | | | SBA Wind Power | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Variable | $\beta$ | $R^2$ | $\Delta R^2$ | $\beta$ | $R^2$ | $\Delta R^2$ | $\beta$ | $R^2$ | $\Delta R^2$ |
| Step 1 | | | | | | | | | | |
| | BK | .48* | .23 | | .49* | .24 | | .39* | .16 | |
| Step 2 | | | | | | | | | | |
| | BK | .82* | .29 | .06 | .81* | .30 | .06 | .67* | .24 | .08 |
| | Number of IDK | .42* | | | .40* | | | .40* | | |

*$p < .01$; SBA = content area learning environment; IDK = I don't know; BK = background knowledge.

## 2.3 Summary

The current study demonstrated that not only was students' understanding related to their background knowledge, but also it was influenced by students' acknowledgements of their lack of background knowledge. Students' utilization of the IDK option explained additional variance of students understanding beyond background knowledge as shown by the hierarchical regression. This suggested that students who selected IDK when answering background knowledge questions had better understanding of the content than those who did not select IDK, but provided more wrong answers instead. This effect was replicated in two comprehension assessments with two groups of high school students. Results of this study were consistent with prior research showing the benefits of including the IDK option in tests (e.g. Wakabayashi & Guskin, 2010). While the precise mechanism for why more IDK selections was associated with higher comprehension (after controlling for BK) cannot be determined from this study, one reason could be that students with a higher incidence of IDK selections are more open to understanding new content than students who proceeded with incorrect knowledge. Students who have incorrect knowledge may have misconceptions that impair their ability to understand what

they read (Meir et al., 2005; € Ozg€ur, 2013). For example, Kendeouand van den Broek (2005) established that students who had previous misconceptions recalled less information after reading, generated more invalid inferences and fewer valid inferences than students with no prior misconceptions. Thus, students who selected wrong answers to BK questions may have had incorrect knowledge that interfered with or “closed the door” to future understanding on that topic. In contrast, students who acknowledge their low knowledge state (high use of IDK) may adjust their standard of coherence (van den Broek, Young, Tzeng, & Linderholm, 1999) and set goals (van den Broek et al., 2001) before they read. One such goal could be to learn new information; in particular for items in which they previously indicated they did not know. If this is correct, then higher use of the IDK option should be associated with learning the previously unknown concepts after they are given a chance to read sources that contain answers to the BK questions. Consequently, the purpose of Study 2 was to explore whether students who were more likely to acknowledge they don’t know, would also be more likely to learn more from reading the materials than students who have incorrect background knowledge. In Study 2, we first replicate the results of Study 1 with a different topic and content-area unit designed for a different age group (middle school). Then, we explore whether the IDK affordance would also result in more knowledge gains (i.e., learning) after reading sources that contained answerers to the BK items.

## 3 Study 2

The aims of Study 2 were two fold. First, we extend and elaborate on Study 1, by evaluating whether we can replicate H1 and H2 (see “Hypothesis” section) with a different BK measure (sentence verification, vocabulary in context) for a different topic (wind power) and a different age group (middle school). Second, we were interested in determining whether student's use of

the IDK option would be related not only to comprehension, but also their ability to learn new information (H3).

## 3.1 Method

### *3.1.1 Participants*

A total of 1,229 sixth and seventh grade students from urban and semi-urban communities in the Northeastern USA participated in the study. All data were collected according to the research institution’s internal review board policy. Due to agreements with the schools, no individual demographic information other than grade level was available to the researchers. However, because the study was part of a large intervention project (LaRusso et al., 2016), we report the demographic information for it, which included 2,933 students in 125 classrooms: gender 51% female; race 40% Black, 28% White, 3% Asian, 27% Latino, 1% Native American/Pacific Islander and 1% other; 83% of students were eligible for free/reduced-price lunch; 8% were English language learners.

### *3.1.2 Instruments*

The SBA used in this study was on the topic of Wind Power and was designed for use with middle school students. The unit consisted of a series of tasks that asked students to “build” a website about wind power. Similar to SBA Ecology and SBA Immigration units, the Wind Power unit also included a wide range of item types. Items required students to identify relevant questions, complete graphic organizers, interpret a chart, evaluate web-search results, make inferences, identify facts and opinions, detect errors and irrelevant information, repair incorrect information, and perspective taking (see Appendix A). The two primary content passages in the assessment were 339 and 450 words in length and had grade level of Flesch Kincaid Grade

level of 9.0 and 8.4 and a Flesch Reading ease of 54.8 and 59 respectively. The TextEvaluatorVR grade level for the two passages was grade 6 and grade 8, respectively. Reliability (Cronbach’s alpha) of the 33-items was a ¼ .88.

#### *3.1.2.1 Background Knowledge*

Two types of tasks comprised the 17-item background knowledge measure, though both were different from the methods used in Study 1. Five vocabulary in context items required students to select the best word that completes the given sentence. Twelve content items were in a True/False format and measured concepts, facts, principles and causal mechanisms described in the passage. Both vocabulary in context and content question types were presented to the students at the start of the SBA before they read the sources (pretest), and again after they read sources that contained answers to these items (post-test). When presented before the sources, these items measured students’ prior BK; when presented after the sources, these items measured students’ learning after having taken the unit. The reliability of the 17-item knowledge pretest section and the corresponding 17-item post-test section was a .57 and a.64, respectively. The relatively lower reliability of the knowledge section is somewhat influenced by the lower total number of items (17 in Study 2, vs. >50 items in Study 1). There were no floor or ceiling effects evident in the distribution of the students’ scores on the various measures, an indication that the SBA was age appropriate for the participating middle school students. The BK measure and the number of IDK responses were calculated in the same way as Study 1. BK score was calculated as the number of correct responses divided by the total number of BK items (n=17).

### ***3.1.3 Procedure***

Data collection followed the same procedures as Study 1. The SBA was web-administered to students during regular 50-minute class period as part of a larger study with collaborating

partners working in the schools. The knowledge task was part of the unit, framed as a glossary-building activity for the wind power website, before students read any of the relevant sources during the comprehension sections of the unit. After students completed the SBA learning unit, they were presented the knowledge task again to evaluate their knowledge gain. The vocabulary in context items appeared first, followed by the content knowledge items. Students again could select IDK if they were not sure of the answers to the knowledge questions. Similar to Study 1, students were told that their performance on the knowledge items would not count towards their final score. However, they were not informed that these items might appear again at posttest.

### 3.2 Results

Table 3 presents the means, standard deviations, and reliability estimates for each of the measures. The correlation between pretest BK and understanding was $r(1205)= .39$, $p < .01$, thus replicating Study 1 results which confirmed H1. In addition, the correlation between pretest BK and posttest BK was $r(1205)=.44$, $p < .01$; and the correlation between understanding and posttest BK was $r(1205)$ ¼ .66, $p < .01$. To replicate Study 1 results regarding H2, we again performed hierarchical regressions to determine whether students’ utilization of the IDK option predicted unique variance in students’ understanding over and above the background knowledge. Results are presented in Table 2 under the “SBA Wind Power” columns and the pattern of results replicates that of Study 1. Specifically, the use of the IDK option explained an additional 8% of the variance in students’ understanding than the effects of background knowledge alone. In addition to replicating Study 1, the main goal of Study 2 is to test H3, that students who are more likely to use the IDK option will have more learning gain than students who select incorrect options on the knowledge test. To test the hypothesis, we performed another hierarchical regression to predict students’ performance on the posttest BK section. Results again showed that

while students’ pretest BK significantly predicted posttest BK, b=.44, p<.01, $R^2$= .19, once controlling for the effect of pretest BK (b=.68, p<. 01) the number of IDK selections positively predicted posttest BK, b=.35, p<.01, and $R^2$ increased to .26. Thus, the number of IDK selections explained another 7% of the variance of posttest BK beyond the effect of pretest BK. In other words, students who selected more IDK options showed more learning gain from the SBA unit. To further examine this effect, we divided students into high vs. low IDK groups based on how much they acknowledged their lack of prior knowledge (use of the IDK option). Students with more IDK responses than incorrect responses in the pretest BK section were classified into a high- IDK category (N=391). Otherwise, students were classified into the low-IDK category (N=816). On average, students in the high IDK group had six IDK responses (SD=2) and three incorrect responses (SD=1); in contrast, students in the low IDK group had two IDK responses (SD=2) and six incorrect responses (SD=2). We also classified pretest and posttest BK performance into three levels. On both measures, performance above the 70th percentile were classified as high while performance below the 30th percentile were classified as low, with middle in between. Table 4 shows distribution of students’ understanding before and after reading the sources. To investigate how the IDK affordance helps learning, students with high background knowledge (top 30%) were left out, because they already knew the answers to most of the items, so learning was unnecessary. Students with low (bottom 30%) or median (middle 40%) knowledge were divided into four groups (students whose score on the posttest BK items was either better or worse than BK items, by high or low use of the IDK response). The chi-square (v2) statistic was used to test whether the use of the IDK option is independent from students’ learning. An odds ratio was also calculated to show the degree of dependence between IDK usage and learning. The v2 test for Table 5 is significant, v2

(1)=58.73, p < .01, suggesting that learning and use of IDK option are not independent (i.e., compensation effect). Students with high incidence of using the IDK option are 2.86 times more likely to show improvement or significant improvement after reading than peers who have comparable amount of knowledge.

**TABLE 3** Means, Standard Deviations and Reliability Estimates (Cronbach's alpha) for the Measures Used in Study 2

| Measure | Mean | Standard Deviation | Cronbach's alpha |
|---|---|---|---|
| Background knowledge before SBA | .53 | .16 | .57 |
| Number of IDK Responses | 3 | 3 | NA |
| SBA Wind Power | .58 | .22 | .88 |
| Background knowledge after SBA | .66 | .17 | .64 |

**TABLE 4** Performance Level Change

| | Posttest BK Items ($n = 17$) | | |
|---|---|---|---|
| Pretest BK items (n = 17) | Low 30% | Mid 40% | High 30% |
| Low 30% | 169 | 124 | 49 |
| Mid 40% | 163 | 208 | 245 |
| High 30% | 19 | 71 | 159 |

**TABLE 5** Four Groups of Students with Low or Median Background Knowledge

| | High IDK | Low IDK | Total $n$ |
|---|---|---|---|
| Students with Improvement | 60% ($n = 207$) | 34% ($n = 211$) | 418 |
| Students without Improvement | 40% ($n = 138$) | 66% ($n = 402$) | 540 |
| Total $n$ | 345 | 613 | 958 |

*Note.* The total number of students is 958 instead of 1207, because 249 students with high background knowledge were excluded from the analyses.

## 4 General Discussion

The purpose of this study was to explore the impact of students’ background knowledge on their ability to understand and learn content. While decades of research has established a connection between students’ understanding and background knowledge, few studies have investigated the effects in a modern reading context, that may demand the purposeful deployment of a set of integrated skills (Britt et al., 2017; Goldman, Britt, et al., 2016; LaRusso et al., 2016; Leu et al., 2013; NGA & CCSSO, 2010). To this aim, we had students engage in a goal directed scenario in which they were asked to understand and learn thematically related text content. We measured students’ background knowledge prior to engaging with the text content, as well as their understanding (Studies 1 and 2) of that content, and their learning of the answers to background knowledge questions (Study 2). Consistent with prior research, we found that students’ background knowledge can facilitate their understanding when the knowledge is correct (Ozuru et al., 2007). Importantly, this study extends prior research by demonstrating the effect of background knowledge when students are engaged in a more complex, goal-driven environment that demands not only basic understanding of text content, but reasoning and applying their acquired knowledge. This facilitation of background knowledge occurred despite the fact that the sequence of activities in the SBA was designed to somewhat reduce the effect of background knowledge by progressively building up students’ understanding over the course of the session via engaging students in the use of learning strategies found effective in supporting student understanding and learning from texts (McNamara, 2007). In one sense, this result is somewhat disappointing, as one might hope that engaging in a sequence of learning strategy activities would be sufficient to support understanding and learning of text content. However, it is consistent with other research that took even larger steps to reduce of the influence of

background knowledge. For instance, Shapiro (2004) found an effect of background knowledge on comprehension even when the content of the passages were fictitious. In short, background knowledge may have an effect on students’ ability to understand content, event when steps are taken to reduce its effect. Background knowledge can be correct or incorrect (Kendeou & O’Brien, 2016), and students should be aware of their lack of knowledge (Tobias & Everson, 2009). In the age of technology and the internet, we are bombarded with a wide range of sources that may provide false information. To compound matters, prior research has shown that incorrect knowledge and misconceptions are difficult to change (Meir et al., 2005; € Ozg€ur, 2013) or fully overcome (Lassonde et al., 2016). Thus, incorrect knowledge may have far reaching consequences for understanding, learning, and one’s ability to make informed decisions. To help investigate some of these issues, we added the use of an IDK option to a background knowledge measure. Such an option could be useful for distinguishing the students who acknowledge their lack of knowledge, from students who appear to have incorrect knowledge. In everyday settings students need to understand and learn text content for which they have little or no knowledge. To the extent that students are able to recognize their knowledge deficits, can they set goals and take strategic actions to help them learn new material? This was the question of the current investigation - to investigate how students’ use of the IDK option impacted their subsequent ability to comprehend and learn new information. We hypothesized that students who acknowledged their lack of background knowledge would be more likely to comprehend and learn new information than students who showed incorrect knowledge. The data from the studies provided some support for this hypothesis. Students who were more likely to acknowledge their lack of background knowledge understood more (Studies 1 and 2) and learned more content (Study 2) than students who selected incorrect BK options.

Students who had a high incidence of IDK responses were nearly three times as likely to learn when they read sources that provided the correct information than peers who had comparable amount of knowledge (Study 2). Thus, some lower knowledge students seemed to partially compensate for their knowledge deficiency, and this was associated with a tendency for these students to acknowledge their lack of knowledge related to the topic. One possible explanation for this effect is that students with higher use of the IDK option are more likely to set goals for learning information than students who selected incorrect knowledge options. When encountering a word or concept they don’t know, they may adjust their standard of coherence (van den Broek et al., 2001) and set a goal for learning new information. In other words, their “don’t know” during the BK task response may cue them to areas of the text where the concepts are covered and subsequently, allocate attention and resources to learn what they previously didn’t know. More generally however, we conclude that in order to better understand the relations between knowledge and comprehension, researchers and educators should consider both students’ BK and students’ perceived lack of BK. As discussed later, measuring only one aspect may overlook key distinctions that might be useful for instruction. Collectively, these results are consistent with the literature on the use of “don’t know” and “unsure” response formats (Scoboria & Fisico, 2013; Wakabayashi & Guskin, 2010). Don’t know responses can provide more information about test takers and may improve the quality of the information gathered (Scoboria & Fisico, 2013). When students are unsure of their initial knowledge states they often learn more information at posttest than students who overestimate their pre knowledge states (Wakabayashi & Guskin, 2010). These “teachable moments” may set the stage for goal setting and subsequent learning by engaging students in appropriate behaviors. The results of the current study extend this finding to the domain of content area learning in a modern complex

reading environment. Knowing what you don’t know, is critical in determining what you understand and have learned. In contrast, students who fail to acknowledge they don’t know and select incorrect answers to the knowledge questions might have misconceptions that impair or close the door on future understanding and learning. This is consistent with the research on misconceptions and conceptual change that has demonstrated that students who have prior misconceptions were likely to recall less information after reading, generated more invalid inferences and fewer valid inferences than students with no prior misconceptions (Kendeou & van den Broek, 2005). It is as if students who select the incorrect option on the BK test, either have incorrect BK before the experiment, or they gain incorrect BK from answering the BK items. Either way, our results suggest that the incorrect BK seems to interfere with subsequent understanding and learning of content relative to the students who acknowledge their lack of background knowledge through IDK responses. Thus, as long as students acknowledge their low level of background knowledge, they seem to have opportunities to learn the knowledge (“You don’t know won’t hurt you”); however, students’ incorrect knowledge seems to harm subsequent understanding and learning (“You don’t know you’re wrong”). Of course, students who select incorrect answers on the knowledge test could be guessing, and therefore they have low knowledge, rather than incorrect knowledge. However if these students know they have low In additional to the theoretical and empirical findings, the methods of the present study also add to the literature in several ways. First, background knowledge was measured directly though sentence verification, topical vocabulary, and multiple choice questions, rather than indirectly via knowledge ratings (e.g., Allwood & Granhag, 1996). This direct method provides a more objective measure of knowledge than self-report ratings. Second, the current investigation used a measure of background knowledge that was directly relevant to the topic of the texts (topical

knowledge), rather than a more general background knowledge measure (e.g., Nelson & Narens, 1980), or a measure of background knowledge that is related to the general discipline (e.g., general science measure in O’Reilly & McNamara, 2007). A more specific measure of topical knowledge is likely to be more relevant to the topic of the texts and more sensitive to learning and comprehension effects. Our findings also generalized across three topics, multiple measures of BK and two grade bands of students. Third, the current experiment (Study 2) investigated students’ ability to repair (or not repair) their incorrect knowledge when given the opportunity in a learning environment that demanded the orchestration of a wide range to complex skills. Because there were no special instructions that told students they should explicitly repair or change their answers in light of the text, any observed changes in student behavior was learner-driven. The current design extends this work by providing a way to distinguish between students who know they have low knowledge, but can learn this information when given the opportunity, from those students who have incorrect knowledge (or think they know), but fail to repair their incorrect knowledge when given the opportunity. By providing a learning environment that can distinguish between these two different student profiles, teachers can supply targeted instruction to help address specific profile needs.

### 4.1 Implications for Instruction

The results of this study and related research have potential applications in education. Administering a background knowledge measure that is relevant to a particular text might benefit instruction. If the background knowledge score indicates that some students have little or no knowledge about the topic, teachers can provide definitions of key vocabulary and provide some factual information and context to facilitate understanding. Importantly, as the results of this study also indicate, background knowledge is not the only construct to consider. Students’

acknowledgement of their level of knowledge is important in determining how much they will learn. Thus, incorporating an IDK option when measuring background knowledge might also provide useful student information for teachers. An IDK manipulation, such as the one included in the current study, can be useful for distinguishing students who know that “they don’t know” from those who don’t know that “they are wrong”. Such a distinction will facilitate instruction because it enables teachers to identify students who might need help with making more accurate knowledge judgments, as well as helping identify students who may have incorrect knowledge that may interfere with learning. But knowing what you don’t know is not enough. Students also need to set appropriate learning goals and to take strategic actions to acquire new facts and conceptual models. Subsequent strategic actions are necessary to repair gaps in one’s understanding. Translating this idea into a concrete action might involve teaching students to set clear reading goals that focus on increasing the student’s standard of coherence. Instilling a mindset for evaluating what one knows and to encourage students to seek resources to learn what they don’t know might be a first step. That is, both background knowledge and an awareness of the lack of knowledge are important for comprehension and learning. For students who have incorrect knowledge and are not aware of it, measures can be taken to challenge their misconceptions through refutation texts and such approaches have been successful in reducing some misconceptions (Broughton et al., 2010; Diakidoy et al., 2016; Lassonde et al., 2016; van Loon et al., 2015).

### 4.2 Limitations and Implications for Future Research

While the results from this study are encouraging, there are a number of issues to address in future research. For instance, the overall reliability of the individual knowledge measures was

less than desirable in Study 2. In ongoing research, we have been working on ways to improve the reliability of the measure by using large numbers of topical vocabulary items (similar to the measures used in Study 1). Second, while the sample size is adequate, the lack of individual level demographic and background information limits any claims about differences in gender and ethnicity that may impact the results. Future experiments should include individual level demographic variables to estimate any individual or group differences that may mediate or moderate the effects. Also, while the experiments included three different types of comprehension tasks (and multiple BK measures), future work should examine whether the results hold for different comprehension tasks and topics. This would include using both content area learning environments and more traditional reading comprehension tests. Third, even though the students in both studies were given the goal or objective of understanding and learning text content as part of the session scenario, they also understood that they were being assessed, not actually being held accountable for learning the content. In the future, we would like to manipulate the context of the study, such that students were instructed more specifically to learn the content as a routine part of their classroom curriculum activities. The current study examined the impact of IDK option on students’ comprehension and learning. Future work should explore the impact of the alternative wordings such as “not sure” option to determine if the same results hold. Building on this idea, future work could also explore whether including both “not sure” and “I don’t know” options simultaneously would impact the results. One limitation of the current work is that there is no definitive evidence to determine whether the students who use the “I don’t know” response are truly low knowledge, or whether they have moderate knowledge, but are less likely to risk an incorrect guess. Having the “not sure” option could help clarify this issue. In addition, future work should examine how influential models of

conceptual change (Kendeou & O’Brien, 2014) and models of learning and expertise (Alexander, 2003) apply to the current study results. For instance, the authors of KReC model (Kendeou & O’Brien, 2014) might assert that increasing the salience and interconnectedness of casual factors of refutation texts should increase the probability that a student will overcome a misconception. Similarly, the Model of Domain Learning (Alexander, 2003) asserts the important developmental relationships between knowledge, interest, strategies and expertise. Unfortunately our study did not vary the causal factors or measure students’ interest, but those are important goals to explore in future work. Finally, while the results of Study 2 lend some support to the notion that the use of the IDK affordance is related to their subsequent learning, the results entertain other possibilities. For example, while 60% of high IDK students showed improvement, 34% of low IDK students also showed improvement (Table 5). Similarly, while 66% of low IDK students did not show improvement, 40% of high IDK students did not improve. Thus, while the use of the IDK affordance is useful for determining which students might improve, there are other factors that were responsible for the gains. Although the present study was not designed to tease out these alternative explanations, it is possible that some of the students who were classified as high IDK were less motivated and thus they selected IDK rather than to answer the question (the 40% group in Table 5). It is also possible that some students who did improve but were classified as low IDK (34% group in Table 5) could have used other successful strategies to learn the new material. Future research should employ designs that can help identify why some readers learn and while others do not.

## 5 Conclusion

In these studies, we have demonstrated the potential for measuring background knowledge with the IDK option as a support to students understanding and learning of text content, in a goal-

directed, complex scenario-based learning context. Judgments such as IDK responses can be an efficient way to separate students who acknowledge their lack of background knowledge, from students who might have incorrect knowledge or misconceptions, which has consequences for understanding and learning. Together, innovative techniques such as these suggest a pathway towards generating more valid inferences and deriving greater instructional utility from content area learning environments, and toward more effective, targeted instructional practices to help students learn from texts.

References

Alexander, P. A. (2003). The development of expertise: The journey from acclimation to proficiency. *Educational Researcher. 32*, 10–14. doi:10.3102/0013189x032008010.

Alexander, P. A. (2012). Reading into the Future: Competence for the 21st Century. *Educational Psychologist*, *47*(4), 259-280

Alexander, P. A., Kulikowich, J. M., & Schulze, S. K. (1994). How subject-matter knowledge affects recall and interest. *American Educational Research Journal*, *31*(2), 313-337. doi:10.2307/1163312

Alexander, P., Sperl, C., Buehl, M., & Chiu, S. (2004). Modeling domain learning: profiles from the field of special education. *Journal of Educational Psychology, 96*, 545-557.

Allwood, C., & Granhag, P. (1996). Considering the knowledge you have: Effects on realism in confidence judgments. *The European Journal of Cognitive Psychology, 8,* 235-256.

Beck, I., McKeown, M., & Gromoll, E. (1989). Learning from social studies texts. *Cognition and Instruction, 6,* 99-158. https://doi.org/10.1207/s1532690xci0602_1

Bråten, I., Johansen, R., & Strømsø, H. I. (2017). Effects of different ways of introducing a reading task on intrinsic motivation and comprehension. *Journal Of Research In Reading*, *40*(1), 17-36. doi:10.1111/1467-9817.12053

Britt, A., Rouet, J. F., & Durik, A. M. (2017). *Literacy beyond text comprehension: A theory of purposeful reading*. New York, NY: Routledge.

Broughton, S. H., Sinatra, G. M., & Reynolds, R. E. (2010). The nature of the refutation effect: An investigation of attention allocation. *Journal of Educational Research, 103,* 407– 423. doi:10.1080/ 00220670903383101

Courtenay, B. C., & Weidman, C. (1985). The Effects of a "Don't Know" Response on Palmore's Facts on Aging Quizzes. *Gerontologist, 25*(2), 177-81.

Cromley, J., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology, 99,* 311-325.

Diakidoy, I. N., Mouskounti, T., Fella, A., & Ioannides, C. (2016). Comprehension processes and outcomes with refutation and expository texts and their contribution to learning. *Learning And Instruction, 41*60-69. doi:10.1016/j.learninstruc.2015.10.002

Dochy, F., Segers, M., & Buehl, M. (1999). The relation between assessment practices and outcomes of studies: The case of research on prior knowledge. *Review of educational Research, 69,* 145-186.

Fancsali, C., Abe, Y., Pyatigorsky, M., Ortiz, L., Chan, V., Saltares, E., ... & Jaciw, A. (2015). The Impact of the Reading Apprenticeship Improving Secondary Education (RAISE) Project on Academic Literacy in High School: A Report of a Randomized Experiment in Pennsylvania and California Schools. Research Report. *Empirical Education Inc.*

Fincher-Kiefer, R., Post, T., Greene, T., & Voss, J. (1988). On the role of prior knowledge and task demands in the processing of text. *Journal of Memory and Language, 27,* 416-428.

Franzke, M., Kintsch, E., Caccamise, D., Johnson, N., & Dooley S. (2005). Summary Street®: Computer support for comprehension and writing. *Journal of Educational Computing Research, 33,* 53–80.

Goldman, S. R., Britt, M. A., Brown, W., Cribb, G., George, M., Greenleaf, C., & ... Shanahan, C. (2016). Disciplinary literacies and learning to read for understanding: A conceptual framework for disciplinary literacy. *Educational Psychologist, 51*(2), 219-246. doi:10.1080/00461520.2016.1168741

Goldman, S. R., Snow, C., & Vaughn, S. (2016). Common themes in teaching reading for understanding: Lessons from three projects. *Journal of Adolescent & Adult Literacy, 60*(3), 255-264. doi:10.1002/jaal.586

Hacker, D. J, Dunlosky, J., & Graesser, A. C. (2009). *Handbook of metacognition in education*. Mahwah, NJ:  Erlbaum.

Hacker, D. J, Dunlosky, J., & Graesser, A. C. (2009). *Handbook of metacognition in education*. Mahwah, NJ: Erlbaum.

Hambrick, D. Z., & Engle, R. W. (2002). Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis. *Cognitive Psychology, 44*, 339-387.

Kendeou, P. & O’Brien, E. J. (2014). The knowledge revision component (KReC) framework: Processes and mechanisms. In D. Rapp, & J. Braasch (Eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences*. Cambridge, MA: MIT.

Kendeou, P., & O’Brien, E. J. (2016). Prior Knowledge: Acquisition and Revision. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reader, text, and context* (pp. 151-163). New York, NY: Routledge.

Kendeou, P. & van den Broek, P. (2005). The effects of readers’ misconceptions on comprehension of scientific text. *Journal of Educational Psychology, 97*, 235-245. http://dx.doi.org/10.1037/0022-0663.97.2.235

Kincaid J. P., Fishburne R. P., Rogers R. L., & Chissom B. S. (1975). *Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel*. (Research Branch Report 8-75). Memphis, TN: Naval Technical Training, U. S. Naval Air Station.

King, A. (1995). Inquiring minds really do want to know: using questioning to teach critical thinking. *Teaching of Psychology, 22*, 13-17.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* Cambridge, UK: Cambridge University Press.

Kintsch, W. (2012). Psychological models of reading comprehension and their implications for assessment. In J.P. Sabatini, E. Albro, & T. O’Reilly (Eds.), *Measuring up: Advances in how we assess reading ability* (pp. 21-38). Lanham, MD: Rowman & Littlefield Education.

Kucer, S. B. (2011). Going beyond the author: What retellings tell us about comprehending narrative and expository texts. *Literacy, 45*, 62-69.

LaRusso, M., Kim, H. Y., Selman, R., Uccelli, P., Dawson, T., Jones, S., ... & Snow, C. (2016). Contributions of academic language, perspective taking, and complex reasoning to deep reading comprehension. *Journal of Research on Educational Effectiveness*, *9*(2), 201-222. https://doi.org/10.1080/19345747.2015.1116035

Lassonde, K. A., Kendeou, P., & O’Brien, E. J. (2016). Refutation texts: Overcoming psychology misconceptions that are resistant to change. *Scholarship of Teaching and Learning in Psychology, 2*(1), 62–74.

Leu, D., Kinzer, C., Coiro, J., Castek, J. & Henry L. (2013). New literacies: A dual-level theory of the Changing Nature of Literacy, Instruction, and Assessment. In D.E. Alvermann, N.J. Unrau, & R.B. Ruddell (Eds.), *Theoretical models and processes of reading* (pp. 1150-1181). Newark, DE: International Reading Association.

Magliano, J.P., McCrudden, M.T., Rouet, J.F., & Sabatini, J. (2018). The modern reader: Should changes to how we read affect research and theory? In M. Schober, M.A. Britt, & D. Rapp (Eds.), *Handbook of discourse processes* (2nd Ed.). New York: Routledge.

Marsh, E. J., Meade, M. L., & Roediger, H. L. (2003). Learning facts from fiction. *Journal of Memory and Language*, *49*, 519-536.

McCrudden, M. T., Magliano, J. P., & Schraw, G. (2010). Exploring how relevance instructions affect personal reading intentions, reading goals and text processing: A mixed methods study. *Contemporary Educational Psychology*, *35*, 229-241.

McCrudden, M. T., Magliano, J. P., & Schraw, G. (Eds.). (2011). *Text relevance and learning from text.* Greenwich, CT: Information Age Publishing.

McDaniel, M., Anderson, J., Derbish, M. H., & Morrisette, N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology, 19,* 494–513.

McNamara, D. S. (1997). Comprehension skill: A knowledge-based account. In M.G. Shafto & P. Langley (Eds.), *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society* (pp. 508-513). Hillsdale, NJ: Lawrence Erlbaum Associates.

McNamara, D. S. (2001). Reading both high-coherence and low-coherence texts: Effects of text sequence and prior knowledge. *Canadian Journal of Experimental Psychology, 55,* 51-62.

McNamara, D. S. (2007). *Reading comprehension strategies: Theories, interventions, and technologies.* Mahwah, NJ: Erlbaum.

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

Meir, E., Perry, J., Stal, D., Maruca, S., & Klopfer, E. (2005). How Effective Are Simulated Molecular-Level Experiments for Teaching Diffusion and Osmosis?. *Cell Biology Education*, *4*(3), 235-248.

Metzger, M. J. (2007). Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research. *Journal of the Association for Information Science and Technology, 58*(13), 2078-2091.

Meyer, B., & Ray, M.N. (2011). Structure strategy interventions: Increasing reading comprehension of expository text. *International Electronic Journal of Elementary Education, 4(1),* 127-152.

Meyer, B., & Wijekumar, K. (2007). A Web-based tutoring system for the structure strategy: Theoretical background, design, and findings. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theory, interventions, and technologies* (pp. 347–374). Mahwah, NJ: Erlbaum.

Muijtjens, A. M., van Mameren, H., Hoogenboom, R., Evers, J., & van der Vleuten, C. (1999). The effect of a “don't know” option on test scores: number-right and formula scoring compared

Murphy, K., & Alexander, P. (2002). What counts? The predictive powers of subject-matter knowledge, strategic processing, and interest in domain-specific performance. *The Journal of Experimental Education, 70,* 197-214.

National Governors Association (NGA) Center for Best Practices, Council of Chief State School Officers (CCSSO). (2010). *Common Core State Standards for English Language Arts*. National Governors Association Center for Best Practices, Council of Chief State School Officers, Washington, D.C.

Nelson, T. & Narens, L. (1980). Norms of 300 General-Information Questions: Accuracy of Recall, Latency of Recall, and Feeling-of-Knowing Ratings. *Journal Of Verbal Learning And Verbal Behavior, 19,* 338—368.

Organisation for Economic Co-operation and Development (OECD) (2016). *PISA 2018 reading literacy framework.* Paris, France: Author. Retrieved at: https://www.oecd.org/pisa/data/PISA-2018-draft-frameworks.pdf

Ogle, D. (1986). K-W-L: A teaching model that develops active reading of expository text. *Reading Teacher, 39,* 564-570.

O’Reilly, T., & McNamara, D. S. (2007). The impact of science knowledge, reading skill, and reading strategy knowledge on more traditional “high stakes” measures of high school students’ science achievement. American Educational Research Journal, 44(1), 161–196.

O’Reilly, T., & Sabatini, J. (2013). Reading for Understanding: How Performance Moderators and Scenarios Impact Assessment Design (Research Report No. RR-13-31). Princeton, NJ: Educational Testing Service.

O’Reilly, T., Deane, P., & Sabatini, J. (2015). Building and Sharing Knowledge Key Practice: What Do You Know, What Don’t You Know, What Did You Learn? (Research Report No. RR-15-24). Princeton, NJ: Educational Testing Service.

O’Reilly T., Weeks, J., Sabatini, J., Halderman, L., & Steinberg, J. (2014). Designing Reading Comprehension Assessments for Reading Interventions: How a Theoretically Motivated Assessment Can Serve as an Outcome Measure. Educational Psychology Review, 26, (3), 403–424.

Özgür, S. (2013). The persistence of misconceptions about the human blood circulatory system among students in different grade levels. *International Journal of Environmental And Science Education, 8*(2), 255-268. doi:10.12973/ijese.2013.206a

Ozuru, Y., Best, R., Bell, C., Witherspoon, A., & McNamara, D. (2007). Influence of question format and text availability on the assessment of expository text comprehension. *Cognition and Instruction, 25,* 399-438.

Ozuru, Y., Dempsey, K., & McNamara, D. S. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and Instruction, 19, 228-242.*

Pennington, H. R., Pachana, N. A., & Coyle, S. L. (2001). Use of the Facts on Aging Quiz in New Zealand: Validation of questions, performance of a student sample, and effects of a don't know option. *Educational Gerontology, 27*(5), 409-416. doi:10.1080/03601270152053438

Radmacher, S., & Latosi-Sawin, E. (1995). Summary writing: A tool to improve student comprehension and writing in psychology. *Teaching of Psychology, 22*, 113-115.

Ravesloot, C. J., Van der Schaaf, M. F., Muijtjens, A. M., Haaring, C., Kruitwagen, C. J., Beek, F. A., & ... Ten Cate, T. J. (2015). The "Don't Know" Option in Progress Testing. *Advances In Health Sciences Education*, *20*(5), 1325-1338.

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

Rouet, J. F., & Britt, M. A. (2011). Relevance processes in multiple document comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Relevance instructions and goal-focusing in text learning* (pp 19-52). Greenwich, CT: Information Age Publishing.

Sabatini, J., Halderman, L., O’Reilly, T., & Weeks, J. (2016). Assessing Comprehension in Kindergarten through Third Grade. Topics in Language Disorders, 36(4), 334–355.

Sabatini, J., & O’Reilly, T. (2013). Rationale for a New Generation of Reading Comprehension Assessments. In Miller, B., Cutting, L., & P. McCardle (Eds), Unraveling Reading

Comprehension: Behavioral, Neurobiological, and Genetic Components, (pp. 100–111). Baltimore, MD: Brookes Publishing.

Sabatini, J., O’Reilly, T., & Deane, P. (2013). Preliminary Reading Literacy Assessment Framework: Foundation and Rationale for Assessment and System Design. (Research Report No. RR-13-30). Princeton, NJ: Educational Testing Service.

Sabatini, J., O’Reilly, T., Halderman, L., & Bruce, K. (2014). Broadening the scope of reading comprehension using scenario-based assessments: Preliminary finding and challenges. Topics in Cognitive Psychology, 114, 693–723.

Sanderson, P. (1973). The ‘don't know’ option in MCQ examinations. *British Journal of Medical Education, 7,* 25-29.

Schneider, W., Körkel, J., & Weinert, F. E**. (**1989). Domain-specific knowledge and memory performance: a comparison of high- and low aptitude children. *Journal of Educational Psychology, 81,* 306–12.

Scoboria & Fisico (2013). Encouraging and clarifying “don’t know” responses enhances interview quality. *Journal of Experimental Psychology: Applied, 19,* 72-82.

Shapiro, A. M. (2004). How including prior knowledge as a subject variable may change outcomes of learning research. *American Educational Research Journal, 41,* 159–189.

Sheehan, K. M. (2016). *A review of evidence presented in support of three key claims in the validity argument for the TextEvaluator text analysis tool*. (ETS Research Report No. RR-16-12). Princeton, NJ: Educational Testing Service. doi:10.1002/ets2.12100

Thompson, R., & Zamboanga, B. (2004). Academic aptitude and prior knowledge as predictors of student achievement in introduction to psychology. *Journal of Educational Psychology*, 96, 778-784.

Tobias, S. & Everson, H.T. (1996). *Assessing metacognitive knowledge monitoring*. (College Board Report No. 96-01). New York, NY: College Entrance Examination Board.

Tobias, S., & Everson, H. T. (2009). The importance of knowing what you know: A knowledge monitoring framework for studying metacognition in education. In D. J. Hacker, J. Dunlosky, A. C. Graesser, (Eds.), *Handbook of metacognition in education* (pp. 107-127). New York, NY: Routledge.

van den Broek, P., Lorch, R. F., Jr., Linderholm, T., & Gustafson, M. (2001). The effects of readers’ goals on inference generation and memory for texts. *Memory & Cognition, 29*, 1081–1087.

van den Broek, P., Risden, K., & Husebye-Hartman, E. (1995). The role of the reader’s standards of coherence in the generation of inference during reading. In R. F. Lorch, Jr., & E. J. O’Brien (Eds.), *Sources of coherence in text comprehension* (pp. 353–373). Mahwah, NJ: Erlbaum.

van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The landscape model of reading: Inferences and on-line construction of a memory representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 71–98). Mahwah, NJ: Erlbaum.

van Loon, M. H., Dunlosky, J., van Gog, T., van Merriënboer, J. G., & de Bruin, A. H. (2015). Refutations in science texts lead to hypercorrection of misconceptions held with high

confidence. *Contemporary Educational Psychology*, *42*39-48. doi:10.1016/j.cedpsych.2015.04.003

Voss, J., & Silfies, L. (1996). Learning from history text: The interaction of knowledge and comprehension skill with text structure. *Cognition and Instruction, 14,* 45-68.

Wakabayashi, T., & Guskin, K. (2010). The Effect of an ‘‘Unsure’’ Option on Early Childhood Professionals’ Pre- and Post-Training Knowledge Assessments. *American Journal of Evaluation, 31*(4) 486-498.

Youjia H., Woods-Groves, S., Ford, J. W., & Nobles, K. A. (2014). Effects of the Paraphrasing Strategy on Expository Reading Comprehension of Young Adults with Intellectual Disability. *Education & Training In Autism & Developmental Disabilities, 49*, 429-439.

Zimmerman, B.J., & Schunk, D.H. (Eds.). (2001). Self-regulated learning and academic achievement: Theoretical Perspectives (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associate.

# Appendix

## Science Study Group Design: Ecosystems

### Overview

*Purpose:* To assess and promote the reading skills associated with learning from science texts.

*Context and scenario:* You are preparing for an upcoming science test. In order to prepare for the test efficiently, you and some classmates decide to form a study group. You are responsible for helping each other identify key concepts, review and learn from readings, and understand scientific terms and data.

*Domain*: Science; *Topics*: Ecosystems; invasive species; science policy

*Sources*: Ecosystem diagram; passage on general information about invasive species; executive order on invasive species; scientific abstracts; scientific definitions; example ecosystems; executive summary of a federal management plan; field notes; special report on invasive species; data tables; web links.

*Skill foci*: topical vocabulary; background knowledge; understand diagrams; supporting evidence; summarization; understand and apply technical vocabulary; identify key concepts, minor details, incorrect information, and opinions; interpret data; identify relevant web sources; paraphrase; integrate multiple sources; update background knowledge based on textual evidence

### Sections

Section 1: What do we already know? This section is designed to measure various aspects of background knowledge. Two forms of background knowledge are assessed: content focused and topical vocabulary. The general background knowledge is measured through a set of topical vocabulary items. More specific content knowledge is measured with multiple choice items that assess knowledge related to the materials covered in the assessment. A "I don't Know" is added to allow students to acknowledge their lack of prior knowledge.

Section 2: Read and summarize important texts. This section measures test takers' ability to form a coherent model of text. Students are asked to summarize a key source before they are asked more demanding questions later. Guidelines are provided to help clarify the expectations of the task and improve scoring. The intent is to encourage a more global level of understanding in relation to many reading tests that often assess comprehension in a piecemeal fashion. This particular source is general and descriptive in nature. It is designed to provide an overview of the topic before more detailed and demanding texts are presented later.

Section 3: Consider evidence and connect to science policy. This section is designed to measure students' ability to understand the nature of scientific evidence and their ability to apply policy to scientific sources. Students are given a federal policy document that describes the various stages in dealing with invasive species. The student is asked to classify the source presented in section 2 according to the stages discussed in the federal policy document. In addition, students are also asked to classify a series of scientific abstracts and citations. This section is designed to underscore the importance of scientific evidence and the important connections between science and policy.

Section 4: Understand and apply scientific terms. This section is designed to highlight the importance of technical vocabulary in science. While many words that appear in a scientific topic are related, terms have nuanced meanings that often have important consequences for both science and policy. Students are provided with definitions of terms surrounding invasive species and they are asked to apply their understanding of them to several examples. The examples describe situations in which animals interact with a particular environment.

Section 5: Say it in your own words. This section is designed to measure students' ability to paraphrase and determine the meaning of sentences. The first task requires the student to select the option that best preserves the meaning of the target sentence, while another task requires them to generate a paraphrase. The text used in this section is an excerpt summary of an invasive species management plan that elaborates on the stages discussed in section 3.

Section 6: Review scientific data. The purpose of this section is to measure students ability to reason with scientific data and apply what they have learned thus far to a specific ecosystem. This section is broken up into three parts: the first set of items requires students to interpret data; the second set requires students to evaluate the relevance of various web sources; the third set measures students' ability to extract key information from all the sources they have read. In particular, the task requires the student to determine whether a statement is a key concept, minor detail, opinion, or incorrect statement. Collectively, this section measures students' ability to integrate and apply all the information read thus far.

Study Group Design: Immigration

## Overview

*Purpose:* To assess and promote the reading skills associated with learning from historical texts.

*Context and scenario:* You are preparing for an upcoming U.S. History test. In order to prepare for the test efficiently, you and some classmates decide to form a study group. You are responsible for helping each other identify key concepts, review and learn from readings, and organize information.

*Domain*: History; *Topics*: US immigration; Ellis Island; immigration acts

*Sources*: Passages about U.S. immigration and Ellis Island and other relevant historical materials (timelines, Acts)

*Skill foci*: topical vocabulary; background knowledge; identify relevant sources; distinguish primary and secondary sources; summarization; understand claims and evidence; identify key concepts, minor details, incorrect information, and opinions; paraphrase; integrate multiple sources; and apply sources to hypothetical situations.

## Sections

Section 1: what do we already know? This section is designed to measure various aspects of background knowledge. Two forms of background knowledge are assessed: content focused and topical vocabulary. The general background knowledge is measured through a set of topical vocabulary items. More specific content knowledge is measured with multiple choice items that assess knowledge related to the materials covered in the assessment. An "I don't Know" is added to allow students to acknowledge their lack of prior knowledge.

Section 2: read and summarize important texts. This section measures students' ability to form a coherent model of text. Students are asked to summarize a key source before they are asked more demanding questions later. Guidelines are provided to help clarify the expectations of the task and improve scoring. The intent is to encourage a more global level of understanding in relation to many reading tests that often assess comprehension in a piecemeal fashion.

Section 3: think about historical evidence and events. This section is designed to assess elements of historical and critical thinking; in particular claim-evidence relations. Students are asked to demonstrate their understanding of various claims made by the author and then are asked to determine which primary source provides evidence for the particular claim. This section also introduces new sources that augment the key source summarized earlier. These new sources outline key laws and timelines of earlier historical events related to immigration. The student is asked to demonstrate their ability to integrate and apply this knowledge to hypothetical situations.

Section 4: decide what is important and what is not. This section is designed to demonstrate students' ability to extract key information from all the sources they have read. In particular, the task requires the student to determine whether a statement is a key concept, minor detail, opinion, or incorrect statement.

Section 5: say it in your own words. This section provides the second half of the key source read in section 2. It is designed to measure students' ability to paraphrase and determine the meaning of sentences. The first task requires the student to select the option that best preserves the meaning of the target sentence, while another task requires them to generate a paraphrase.

Section 6: organize what you know. This section is designed to assess causality and the elements that drive it. Students are provided with a graphic organizer and are then required to fill in the correct sequences of causes and supporting evidence. The task is scaffolded in two ways that make it more manageable. First, some of the cells in the graphic organizer are filled in to help the student know what is expected. Second, the task is broken up into two sections, one for causes and one for supporting evidence. This is designed to both reduce working memory load, and to correct any mistakes students might have had about the causes in part 1.

# Wind Power

## Overview

*Purpose:* To assess and promote the reading skills associated with learning from science texts.

*Context and scenario:* A wind farm has been proposed for your community. Your class has decided to create a Web site about wind power to help members of the community become more informed about the subject.

*Domain*: Science; *Topic*: Wind power

*Sources*: Expository text on wind power; expository text on pros and cons of wind power; community forum; web links; simulated peer responses

*Skill foci*: background knowledge; summarization; identify relevant questions, complete graphic organizers, interpret a chart, evaluate web-search results, make inferences, identify facts and opinions, detect errors and irrelevant information, repair incorrect information, and perspective taking

## Sections

Section 1: Create a glossary. - This section is designed to measure various aspects of background knowledge to help qualify the interpretation of the score. An “I don’t Know” is provided to allow students to acknowledge their low knowledge. The exact background knowledge items appear later in the test after students have read materials that provide answers to these questions (section 5). This design allows us to track whether students learn the previously unknown information when given the opportunity to read about it.

Section 2: Learn about wind power. This section measures students' ability to form a coherent model of text. Students are asked to summarize a key source before they are asked more demanding questions later. Guidelines are provided to help clarify the expectations of the task and improve scoring. The intent is to encourage a more global level of understanding. This particular source is general and descriptive in nature. It is designed to provide an overview of the topic before more detailed and demanding texts are presented later. Students must also chronologically order events in a graphic organizer, use a bar chart to answer additional questions, and fill in a graphic organizer comparing wind turbines and electric fans.

Section 3: Find More information about wind power. This section is designed to measure students' ability to understand evidence and their ability to choose sources. Students are given a list of links just as they would appear in an actual search engine and is asked to pick which website is the best source and explain why other sites are poor sources.

Section 4: The possibilities and challenges of wind power. This section is designed to highlight the conflicting views about wind power and assess the student's ability to form a coherent model from multiple sources. The student is asked to complete graphic organizers that represent the structure of key concepts and ideas as well as identify problems with various forms of alternative energy.

Section 5: See what you've learned. This section is designed to measure the ability to learn from text. The questions in this section were previously asked in the background knowledge section 1. However, at this point in the unit students have read sources that provide answers to these questions. This allows us to track whether students changed their answers in section 1 to update them in light of the information they have read. For students who select incorrect options in

section 1 we can determine if they hold on to their incorrect beliefs, or whether they change them in light of what they read.

Section 6: The community forum. The student enters a forum with comments from different users which must be organized as fact, opinion, incorrect, or off topic. Finally, the students must decide the best post to respond to the comments in the forum, some of which require the student to correct a post, while others require the student to engage in aspects of perspective taking.