A Checklist for Use with the Lexile Readability Formula When Choosing Materials for Writing Center Self-Access Libraries

John R. Baker

Faculty of Foreign Languages, Ton Duc Thang University, Ho Chi Minh City, Vietnam

Biodata

John R. Baker’s research interests include second language writing and reading, self-access and writing center administration, various literature interests, and how these come together in an interdisciplinary nature. Correspondence regarding this article can be sent to the Faculty of Foreign Languages, Ton Duc Thang University, 19 Nguyen Huu Tho, Tan Phong Ward, District 7, Ho Chi Minh City, Vietnam.

Email: drjohnrbaker@tdtu.edu.vn

Abstract

Writing center self-access libraries have a firm place in North American and non-North American contexts, as do the composition-related materials therein, i.e., rhetorics (e.g., anthologies of paragraphs and essays); however, as the literature has historically shown, when choosing texts for library shelves, readability needs to be considered.

Traditionally, the readability of rhetorics has been explored through single-step quantitative approaches that employ readability formulae (e.g., the Lexile readability formula). These formulae have been praised for their ability to effectively measure two variables (i.e., semantic and syntactic difficulty) but have been criticized because they do not explore the many other features that influence text difficulty. To embrace the strengths and address the weaknesses of

---

1This paper was funded by Ton Duc Thang University (TDTU), Ho Chi Minh City, Vietnam. Questions regarding this research can be sent to Dr. John R. Baker, Ton Duc Thang University, 19 Nguyen Huu Tho St, Tan Phong Ward, Dist. 7, Ho Chi Minh City, Vietnam. Email: drjohnrbaker@tdtu.edu.vn.
these formulae, this paper suggests a two-step hybrid method: (a) a quantitative evaluation of
texts and students’ reading levels via Lexile measures and (b) a qualitative exploration of the
features that the Lexile formula does not explore (i.e., via a checklist). To provide a checklist
for this procedure, this sequential mixed methods study explored what features not considered
by the Lexile readability formula affect the reading experiences of students when reading
exemplars from rhetorics. The study found that 14 features beyond the two measured by the
Lexile readability formula influence users’ difficulty with the material. This paper also
discusses the relevance of the study to the field and poses suggestions for future study.

**Keywords:** Writing Center Administration, Lexile, Readability, Rhetorics, Self-access

**Introduction**

In his 1984 article The Idea of the Writing Center, North firmly established the notion that self-
access libraries have a place in writing centers. Since then, this idea has become a historical
constant (Baker, 2018, Childers, 2006; Harris, 1992; Simpson, 1985), as have the composition-
related texts that directors choose to stock their shelves with (Baker, 2019; Kincead & Harris,
1993). A review of the literature also shows that writing centers outside the North American
context (e.g., in Asia), which are few in number (Paiz, 2017; Tan, 2011) but growing (Chang,
2013; Murphey, Falout, Elwood & Hood, 2009), also provide self-access libraries that offer
students an array of English composition-related texts, including rhetorics, which are
rhetorically organized anthologies of paragraphs and essays “which explicate major rhetorical
forms, present sample texts exemplifying major rhetorical patterns, and offer procedures to
show students how to reproduce these patterns and genres in their own writing” (Ferris &
Hedgcock, 2005, p. 130)\(^2\).

Accepting that self-access resource libraries have an important place in writing centers is one
thing; stocking them with materials is another matter (Baker, 2019). And this is where the
challenge lies because, as Baker noted, the process of selecting texts (e.g., rhetorics) for writing
center resource libraries, in both North American and non-North American settings (e.g., in
Asia) is not simply a matter of purchasing texts and placing them on shelves. Instead, it requires

\(^2\) Rhetorics are not to be confused with thematic readers, anthologies which are primarily thematically
organized.
writing center staff to determine which texts are appropriate for the students who may visit the center.

Appropriacy (i.e., having materials at the appropriate reading levels) is essential, for (a) rhetorics provide genre-specific models of academic writing and (b) apprenticing writers who engage genre-specific materials as part of the reading-writing relationship tend to exhibit competency in the use of conventional features when composing their own texts, with as much as a .50 to .70 correlation (Grabe, 2003, 2016). However, such gains cannot be realized unless students can comprehend the reading materials that they encounter. Specifically, students “can’t learn much from books they can’t read” (Allington, 2002, p. 16). To address this problem, writing center staff must undertake a procedure “familiar to all those who choose books for their own use” (Gilliand, 1972, p. 12). They need to select texts that will be a good fit for the readers who might use them (Baker, 2019), the study of which has come to be called readability (Gilliand, 1972).

**Literature Review**

Historically, readability has been defined in many congruent ways, with each definition emphasizing the importance of considering the text, the reader, and the interaction between the two. A review of the many definitions of readability that have been formulated illustrates this point (see Chall, Bissex, Conrad, & Harris-Sharples, 1996; Dale & Chall, 1949; Gilliand, 1972; Harrison, 1980; Kintsch & Miller, 1981; Kintsch & Vipond, 1979; Klare, 1963; Schirmer & Lockman, 2001). One metaphor-based definition that is especially informative is from Chall, Bissex, Conrad, and Harris-Sharples’ (1996). These authors likened readability to an iceberg. Beneath the water level, there are “various sources of difficulty. The more difficult the passage, the greater the ice beneath” (p. 6). This definition elucidates the concept of readability because it indicates that readability, similar to an iceberg, is not one solid, homogeneously transparent entity. Instead, it is a heterogeneous mix of features that constitute what Goodman (1967) called the *psycholinguistic guessing game of reading*—a game that requires a consideration of the many features that contribute to the complex phenomenon known as reading.

A historical meta-analysis (1781 to present) confirms our field’s concern with this multifaceted phenomena. As presented in Table 1, the literature shows that researchers have historically cited one or more features that contribute to readability, but no study has considered all of these features in one work, something that needs to be considered when holistically exploring how readable texts are for the students who may read them.
Table 1. Features Commonly Cited in Readability Literature

1. Vocabulary: The number of unfamiliar, abstract, figurative, or technical words in a text. (1912-2019)

2. Sentence Length: The number of words in each sentence. (1893-2019)
   - Blau, 1982; Coleman & Miller, 1968; Coleman, 1962; Dwaik, 1997; Freedle & Kostin, 1992; Glazer, 1974; Guarino & Perkins, 1986; Guo, 2008; McElree, 2000; McElree, Foraker, & Dyer, 2003; McLaughlin, 1969; Mahmoud, 2015; Mikk, 2008; Mikk & Kukemelk, 2010; Pearson, 1974; Rusek, & Vojíř, 2019; Sherman, 1893; Yoo & Lee, 2017

3. Vocabulary in Context: How well the words or sentences surrounding unfamiliar words help the students to understand them. (1993-2018)
   - Ahmad, Muhammad, Kasim, 2018; Ames, 1966; Bengeleil & Paribakht, 2004; Chern, 1993; Cooper, 1999; Davoudi & Nassaj, 2016; de Bot, Paribakht, & Wesche, 1997; Dubin & Olshtain, 1993; Fraser, 1999; Haynes & Baker, 1993; Haynes, 1993; Laufer, 1997; Nassaj, 2003; Paribakht & Wesche, 1999; Rokni & Niknaqsh, 2013; Shen & Wu, 2009; Shen, 2005; Shokouhi & Askari, 2010

4. Background Knowledge Required: How familiar the students are with the topic of a text. (1781-2019)

5. Interest: How interested the students are in the topic presented in a text. (1895-2019)
   - Ainley, Hidi, & Berndorff, 2002; Asher, 1980; Atamturk, 2018; Baldwin, Peleg-Bruckner, & McClintock, 1985; Bargh & Schul, 1980; Belloni & Jongsma, 1978; Benware & Deci, 1984; Bernstein, 1955; Boscolo & Mason, 2003; Brantmeier, 2006; Bugel & Buunk, 1996; Carrell & Wise, 1998; Chodkiewicz, 2016; Choi, 2006; Entin, 1981; Ergüçtin, 2010; Felkin & Felkin, 1895; Kelsen; 2016; Leloup, 1993; Lin, Zabrucky, & Moore, 1997; Mikk & Kukemelk, 2010; Schiefele &


7. Length: (a) The number of words in a text and (b) the number of words in a paragraph. (1894-2018)


8. Logical Rhetorical Organization: How the ideas are arranged in a text to help them flow logically from one to another. (1975-2017)


9. Structure: How well a text is organized (e.g., an introductory paragraph which contains a clear thesis statement, body paragraphs that contain a topic sentence, supporting details, and a concluding paragraph). (1975-2011)


10. Signal Words: Whether a text contains words that indicate the flow of information (e.g., first, next, finally, etc.). (1917-2018)

Aidinlou & Pandian, 2011; Cooper, 1984; Dechant, 1973; Geva, 1992; Meyer, 1975; Miccinati, 1975; Pauk, 1974; Quan, 2008; Rezaee & Norouzi, 2011; Roen,
11. Punctuation: The use of periods (.), question marks (?), exclamation marks (!), commas (,), colons (:), semicolons (;), dashes/hyphens (-), ellipsis ( . . . ), etc. (1919-2007)

Abbott, 2006; Backscheider, 1972; Carr, 1978; Carver, 1970; Durkee, 1952; Graesser, McNamara, & Louwerse, 2003; Hasbrouck, Ihnot, & Rogers, 1999; Nagy, 2007; Neff, 1932; Shih, 1992; Summey, 1919; Tailfever & Pugh, 1998; Van Diem, 1969; Weaver, Holmes, & Reynolds, 1970

12. Format: The physical appearance of a text (i.e., font, type size, spacing, line length). (1881-2018)

Burt, 1959; Campbell, Marchetti, & Mewhort, 1981; Cohn, 1886; Crosland & Johnson, 1928; Griﬃng & Franz, 1896; Hartley and Mills, 1973; Hartley & Rooum, 1983; Haque, et. al., 2018; Hojjati & Muniandy, 2014; Hvistendahl & Kahl, 1975; Josephson, 2008; Kahl, 1974; Kravutske, 1994; Lonsdale, Dyson, & Reynolds, 2006; Lund, 1999; Moriarty & Scheiner, 1984; Muncer, Gorman, & Bibel, 1986; Paterson & Tinker, 1932 a, b; Pinelli, Glassman, & Cordle, 1982; Pyke, 1926; Robinson, Abbamonte, & Evans, 1971; Roethlein’s, 1912; Schriver, 1997; Snow, 1925; 1999; Stone, Fisher, & Eliot, 1999; Tinker & Paterson, 1929; Tinker, 1963a, b; Van Rossum, 1997; Venezky, 1984; Weber, 1881 (see Pyke, 1926); Wieldon, 1995

13. Advance Organizers: Text based introductions prior to a text. (1960-2012)


14. Adjunct Questions: Pre- and post-questions (i.e., the questions before and/or after a text). (1917-2018)

Anderson & Biddle, 1975; Andre, 1987; Brunn, Weidlich, & Bastiaens, 2018; Caverly, 1926; Distad, 1927; Germane, 1920; Hamaker, 1986; Hausman & Rhodes, 2018; Holmes, 1931; Lin & Chen, 2006; Maxwell, 1921; McElroy, 1934; Rothkopf, 1965, 1966, 1970; Shokouhi & Parvaresh, 2010; Thorndike, 1917; Washburne, 1929

Note: Some researchers have only cited one feature (see Lewis, 1894) and thus are only in one category. Others have cited more than one feature and are therefore in two or more categories (see Freedle & Kostin, 1991).

a A brief (usually one paragraph), abstract prose passage presented in written form prior to the text.
Readability Research with Rhetorics

Although many features have been found to contribute to readability, as illustrated in Table 1, most of the studies that have investigated the appropriacy of rhetorics, albeit they are few in number and dated (Auvenshine, 1978; Cline, 1971; Dunn, 1983; Fox, 1978; Morrison, 1978), did so in native speaker (NS) contexts and explored only two features, namely, semantic (vocabulary) and syntactic (sentence length) difficulty. These studies, and most studies of the era, investigated these features by using hand-calculated readability formulae that focused on only these two features. Such formulae were typically employed because, although other formulae had been produced that included a limited number of features beyond these two, researchers generally concluded that further additions beyond these two variables added relatively little predictive validity compared to the added application time involved (Klare, 1975). As a result, studies have traditionally used a single step, three-stage quantitative process: (a) an assessment of the readability level(s) of the texts via a readability formula; (b) an examination of the students’ reading levels with a standardized reading level test that could be correlated with the formula’s results; and (c) a comparison of the two. And each of these studies have reported a gap between students’ reading ability and the difficulty of the required texts.

Research with Rhetorics and English Language Learners (ELLs)

Work with rhetorics in NS settings is noticeably limited and dated, whereas work with ELLs is more recent but, again, relatively scarce. One recent work was completed by Baker (2019) with ELLs in an Asian Context (i.e., Taiwan). Noting that readability formulae have been validated for use with ELLs (see Greenfield, 2004), Baker employed a more recently developed computer-aided formula, specifically, the widely used commercial Lexile readability formula. Baker also found a mismatch between text readability levels and students’ reading levels. That is, the majority of the exemplars (i.e., the paragraphs and essays) therein were too difficult for the target population, namely, ELL undergraduates who visited the campus writing center. Baker also noted that although his results were helpful, further research needs to be performed. Further research is needed because although readability formulae, such as the Lexile formula used in his study, provide good quantitative estimates of text difficulty via a consideration of two features (i.e., semantic and syntactic features), readability formulae have not been met with universal acceptance (Chall & Dale, 1995; Hiebert, 2011). That is, “they have been praised in that they provide an objective quantitative estimate of two factors which have been found to be good predictors of readability” (Baker, 2019, p. 6). However, “they have drawn criticism for the same reason. That is, they usually focus only on these two factors and thus do not explore
other qualitatively assessable areas of readability which are highly relevant to text adoption decisions.”

Of course, as with early formulae, critics have attempted to address the shortcomings of two-factor formulae by developing other formulae that explore additional features (e.g., Coh-Metrix). However, as with earlier alternative formulae, these formulae have not been widely used for a variety of reasons. Coh-Metrix, for instance, is a helpful formula that offers a wide variety of research opportunities (McNamara, Graesser, McCarthy, & Cai, 2014). However, although it addresses a limited number of additional features (albeit not all of the features listed in Table 1), it (a) has complicated algorithms, (b) does not offer sufficiently large text samples to allow for a complete analysis of typical texts, and (c) it has not been employed by the popular commercial standardized tests that evaluate students’ reading levels as part of the three-stage process. Thus, as the Lexile formula successfully addresses each of these three areas, despite addressing only two features, it remains a widely used, commercial formula for text adoption purposes. As such, a hybrid approach needs to be employed to consider the other features that the Lexile formula does not measure.

**Why Use a Hybrid Approach?**

Such a hybrid approach, although still in its infancy, is promising, as Armbruster (2016) argued. For a hybrid approach, one uses a quantitative readability formula and then a qualitative checklist to explore the features not measured by readability formulae. Zakaluk (in 1985) and later with Samuels (Zakaluk & Samuels, 1988) provided an early example. They suggested two categories: (a) outside-the-head factors (i.e., readability as measured by readability formulae, e.g., the Fry formula and adjunct comprehension aids) and (b) inside-the-head factors (i.e., word recognition skills and knowledge of topic). Drawing on these two categories, Zakaluk suggested the use of a nomograph, a graph that visually sets the two factors alongside one another in a way that they can be plotted to determine the predicted level of comprehension as indicated on a center line.

Chall and Dale (1995) also offered a hybrid approach. Reviewing work on cognitive and organizational factors (Kemper, 1983; Kintsch & Miller, 1981; Meyer, 1982) and describing this work as focusing on *The New Readability*, they presented their 1995 readability formula (i.e., the Dale-Chall new readability formula) and offered two checklists to be used with it. The first checklist is used to judge reader characteristics, including whether readers are likely to be interested in the topic and the way that it is presented. The second checklist is used to determine
the cognitive-structural aspects of a text: (a) the prior knowledge expected of the reader; (b) vocabulary and concepts; (c) overall organization; and (d) the use of headings, questions, illustrations, and physical features of the text.

Other researchers have also supported hybrid approaches (Fry, 2002; Gunning, 2003; Metametrics, 2010; Meyer, 2003; Weaver, 2000). Weaver (2000) explained that a hybrid approach is beneficial because “[u]sing readability formulae and subjective criteria [together] reduces the risk of presenting students with a seemingly appropriate book but one they cannot read due to format, language, structure, or content” (p. 33). Some researchers have been even more optimistic. Chall (1996), for example, predicted that hybrid approaches will be the future of readability.

**Why Use a Hybrid Approach with the Lexile Readability Formula?**

The readability formula chosen for this study (i.e., the Lexile readability formula) has received similar attention. Gunning (2003), for example, suggested that using a hybrid approach is prudent, as “although teachers might use Lexiles…, they need to go beyond the numbers… [and] complement the objective data yielded by the formula with subjective judgment” (pp. 182-186). Metametrics, the company behind the Lexile readability formula, also concurred that such an approach has merit:

> A Lexile measure is based on two strong predictors of how difficult a text is to comprehend: word frequency and sentence length. Many other factors affect the relationship between a reader and a book…. The Lexile measure is a good starting point in the book-selection process, but these other factors should be considered when making a decision about which book to choose. (Lexile, 2010)

In response, a variety of teacher preparation texts have offered very general checklists that could theoretically be paired with readability formulae (generally for the selection of textbooks in the content areas) (see Armbruster, 2016), and some of these texts have even suggested that their checklists be used with the Lexile formula (see Lapp, Moss, Grant, & Johnson, 2015). These checklists are indeed valuable, as they list a limited number of features in the form of general guides that could be used to consider content area textbooks. However, these checklists (a) explore only a limited number of features, (b) are based on good practice, not empirical examination, and (c) are not specific to the genre of rhetorics (and the exemplars found therein, i.e., paragraphs and essays).
The Research Gap That Needs to Be Addressed

Although (a) quantitative studies that use readability formulae give a preliminary quantitative picture of the difficulties that postsecondary ELLs have with rhetorics, (b) the virtue of a hybrid approach that uses the Lexile formula with checklists has been recognized, (c) and a small number of checklists (based on good practice, not empirical study) to select content area textbooks have been suggested that focus on a limited number of features, a gap remains. That is, a comprehensive empirical exploration of what features beyond those considered by the Lexile readability formula influence students’ perceptions of reading difficulty when reading exemplars taken from rhetorics has yet to be conducted and included in a genre-specific checklist when exploring rhetorics with a hybrid approach. This study intends to fill this gap.

To address this gap, this article explores a question posed by Baker (2019): Which factors beyond those measured by the Lexile readability formula influence students’ perceptions of reading difficulty when reading exemplars taken from rhetorics?

Methods

To collect and analyze the data needed to answer the research question, Creswell’s (2018) sequential mixed methods research design was adapted and used in this study. This research design is illustrated in Figure 1.

<table>
<thead>
<tr>
<th>Step 1</th>
<th>Step 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Quantitative</strong>&lt;br&gt;An exploration of which features beyond those measured by the Lexile readability formula students perceive as contributing to text difficulty when reading exemplars taken from the rhetorics (i.e., paragraphs and essays) via a cline-questionnaire procedure.</td>
<td><strong>Qualitative</strong>&lt;br&gt;A further exploration of the results of the cline-questionnaire procedure via semi-structured retrospective interviews to add scope and breadth to the findings.</td>
</tr>
</tbody>
</table>

Figure 1. Sequential Research Design.

According to the sequential design illustrated in Figure 1, a bilingual research assistant and I (hereafter referred to as “we”) administered an untimed two-stage process that included (a) a
quantitative cline (difficulty ranking) and questionnaire procedure and (b) qualitative semi-structured retrospective interviews.

**Description of the Setting**

We conducted the two-stage study at Jinwen University of Science and Technology in New Taipei City, Taiwan. The university maintains a writing center that serves the entire student community (9,000 undergraduate and graduate students enrolled in one of 16 majors) and is staffed by paid student-tutors and a small number of dedicated teacher volunteers.

**Description of the Participants**

Consistent with the nature of qualitative theory, we purposively selected the informants to best help us understand the problem (Creswell & Creswell, 2018). In accordance with Kvale’s (1996) suggestion that the number of informants tends to be 15 ± 10 in interview studies, we identified a smaller cluster sample (n = 14) (Table 2) from a larger sample (N = 91), which included individuals who had completed the Scholastic Reading Inventory (SRI) exam (a standardized reading assessment that provides Lexile measures), as reported during the quantitative phase of Baker’s (2019) study.

The informants selected for the cluster sample had received Lexile scores in the top 15% of their class, as provided by the SRI (range 864-928L). Having such high scores indicated that the participants would be able to examine a wide range of exemplars, which would assist us in holistically exploring the research question.

After the 14 informants were identified, we asked them by email if they would be willing to participate in a paid (i.e., 1,000 New Taiwan Dollars, approximately 32 U.S. Dollars, per participant), follow-up, post-course interview. We chose to conduct paid interviews to help ensure that the informants would perform to the best of their abilities. Twelve informants consented and were provided with pseudonyms (alphabetized from A to O) to protect their anonymity.

**Table 2. Characteristics of the Respondents**

<table>
<thead>
<tr>
<th>Gender</th>
<th>Age</th>
<th>Lexile Measures</th>
</tr>
</thead>
</table>
| Kala   | Female | 20  
<p>|        |       | 21  | 864L |</p>
<table>
<thead>
<tr>
<th>Name</th>
<th>Gender</th>
<th>Age</th>
<th>Lexile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jacob</td>
<td>Male</td>
<td>20</td>
<td>869L</td>
</tr>
<tr>
<td>Eve</td>
<td>Female</td>
<td>20</td>
<td>861L</td>
</tr>
<tr>
<td>Marsha</td>
<td>Female</td>
<td>20</td>
<td>877L</td>
</tr>
<tr>
<td>Harold</td>
<td>Male</td>
<td>21</td>
<td>837L</td>
</tr>
<tr>
<td>Linda</td>
<td>Female</td>
<td>20</td>
<td>892L</td>
</tr>
<tr>
<td>Ben</td>
<td>Male</td>
<td>21</td>
<td>858L</td>
</tr>
<tr>
<td>Cara</td>
<td>Female</td>
<td>20</td>
<td>828L</td>
</tr>
<tr>
<td>Dan</td>
<td>Male</td>
<td>20</td>
<td>870L</td>
</tr>
<tr>
<td>Nelson</td>
<td>Male</td>
<td>21</td>
<td>869L</td>
</tr>
<tr>
<td>Olivia</td>
<td>Female</td>
<td>21</td>
<td>926L</td>
</tr>
<tr>
<td>Annie</td>
<td>Female</td>
<td>20</td>
<td>928L</td>
</tr>
</tbody>
</table>

Note: The participants are ordered by Lexile measures.

The informants were undergraduate Taiwanese ELLs seeking degrees in Applied English, seven women (with a mean age of 20.14 years) and five men (with a mean age of 20.8 years). The purposive sampling of students was based on their Lexile measures (range 864-928) rather than their demographic characteristics (e.g., ethnicity, gender, and age).

**Procedure for the Cline (difficulty ranking) and questionnaire.**

**Cline (difficulty ranking) phase**

This stage of the study, the cline (hereafter referred to as “difficulty ranking”) and questionnaire procedure, included two untimed phases: (a) a difficulty ranking and (b) a questionnaire. The purpose of the difficulty ranking phase was to have the participants read five exemplars taken from the rhetorics and place them in order of difficulty so that the participants would be able to reflect on this activity while completing a closed-response questionnaire. For the difficulty ranking phase, the exemplars were presented to the informants in random order, and the ranking criteria were withheld to ensure that the informants engaged in the type of decision-making process that is “normally used when making such judgments” (Chall, Bissex, Conrad, & Sharples 1966, p. 77).
Description of the exemplars

We chose five exemplars for the difficulty ranking phase of the study. We took these exemplars from the rhetorics identified as being available on the local market by Baker (2019). The exemplars are listed in Table 3.

Table 3. Exemplars Chosen for the Study

<table>
<thead>
<tr>
<th>Exemplars</th>
<th>Lexile Measures</th>
</tr>
</thead>
</table>

We also selected the exemplars (N = 12) from those that Baker (2019) identified (N = 867). We purposively chose them to be below, within, and slightly above the informants’ Lexile range. Because of the informants’ low Lexile range (i.e., 828L-928L) and to provide enough essays to encourage the informants to thoughtfully consider their difficulty rankings, we purposively chose the exemplars to be approximately 100L apart to allow the students to reflect on the features that influenced their perceptions of difficulty instead of choosing a higher level (e.g., 200L apart), which would have made the Lexile ranking more obvious but would have reduced the number of exemplars available for ranking.

We acknowledge that the exemplars contained the features listed in Table 1 (each in a different mix), which facilitated the study, but they were not purposively selected for this reason. We avoided purposively selecting exemplars based on their features in order not to load the dice with regard to aiming to achieve expected results.
The Questionnaire Phase

The next phase was the closed-response Likert questionnaire. This phase was conducted to help the informants reflect upon why they ranked the difficulty of the exemplars in the way that they did and describe their ranking processes in such a way that would provide insight into what other factors beyond those measured by the Lexile readability formula that they believed influenced their perceptions of difficulty when reading the exemplars taken from the rhetorics. Once the informants completed the procedure, we ranked the results in frequency and percentage tables.

Description of the Questionnaire

We adapted the questionnaire (Appendix) from the features outlined in Table 1 that have been found to contribute to reading comprehension. It is important to note that we separated one feature described in Table 1—text length (i.e., overall length and paragraph length)—into two questions: (a) overall text length and (b) paragraph length. We also ordered and grouped the questions in a way that we felt would help the informants understand the questionnaire and best reflect on the essays. The question about vocabulary, for example, preceded the question about vocabulary in context.

To ensure the reliability of the questionnaire, we translated the original form into the students’ L1 (i.e., Mandarin) using a back-translation procedure. The translation was then checked by a second translator for accuracy. To further ensure reliability regarding the translation of the questionnaire, we conducted a pretest with a small number of respondents who were not part of the sample used in the study (n = 2). We conducted the pretest to reveal potential ambiguities that may have resulted from the translation of the original questionnaire.

Interview

After ranking the difficulty of the exemplars and completing the questionnaire, the informants engaged in semi-structured retrospective interviews in accordance with Creswell and Creswell (2018). Each interview began with structured questions from the questionnaire that were followed up with semi-structured prompts which later became open-ended (Cohen, Manion & Morrison, 2018). The research assistant (who acted as a translator when necessary) was present to assist with any language difficulties. This procedure was repeated with each of the questions to triangulate the data from the questionnaire.
We also used an observational protocol that included both video and audio recordings of the interviews. The interviews lasted for an average of 32.5 minutes (range 19.3-57.4). Variations in the interview length depended on how much information each informant had to offer and how much translation was required.

**Analysis**

After the interviews were completed, we transcribed the audio recordings, and the informants (with the assistance of the translator when necessary) checked their transcripts. Once these steps were completed, we explored the informants’ responses by using Erlandson, Harris, Skipper, and Allen’s (1993) emergent category analysis procedure. This procedure was used to allow the “categories to follow data rather than precede them” (p. 112). After all of the responses had been coded, “to add strength and fertility to the entire analysis,” we conducted a “second-level group debate” (pp. 128-129). To complete the analysis, we followed previous suggestions that qualitative data can be quantitatively presented in table form (Creswell & Creswell, 2018; Miles & Huberman, 1994), and organized the categories that we had created and presented the results in tables.

**Results**

By using an adaptation of Creswell and Creswell’s (2018) sequential research design, this study explored the features related to readability beyond the two features measured by the Lexile readability formula (i.e., vocabulary and sentence length) that the participants believed to affect the readability of exemplars in rhetorics. This was explored via two data collection techniques: (a) a difficulty ranking and questionnaire procedure and (b) a qualitative semi-structured retrospective interview. We invited fourteen participants, 12 of whom agreed to participate and were given pseudonyms alphabetized from A to O. One participant (O), however, later reported that she was unable to participate, and another (C) did not successfully complete the study; therefore, only 10 participants’ data were considered.

The first stage (i.e., the difficulty ranking and questionnaire procedure) had two untimed phases: (a) a difficulty ranking and (b) a questionnaire. In the difficulty ranking phase, the informants read five exemplars from the rhetorics and ranked them in order of difficulty (easiest to most difficult) so that they could reflect on this activity during the questionnaire phase.
As the purpose of the ranking procedure was to prepare the informants for the next phase (i.e., the questionnaire), the data regarding the informants’ rankings are shown in Table 4. We provide these data to help the reader follow the study, not to draw conclusions about the confidence of the informants’ rankings in relation to one another. We also do not provide these data as a source of validation (or invalidation) of the Lexile measures with this population or rhetorics, as (a) this is not the focus of this study, and (b) the tight grouping of the essays’ Lexile levels (i.e., 100L and less) would not lend itself to such interpretations.

Table 4. The Results of the Informants’ Difficulty Ranking

<table>
<thead>
<tr>
<th></th>
<th>Annie</th>
<th>Ben</th>
<th>Dan</th>
<th>Eve</th>
<th>Harold</th>
<th>Jacob</th>
<th>Kala</th>
<th>Linda</th>
<th>Marsha</th>
<th>Nelson</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easiest</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>d</td>
<td>d</td>
<td>b</td>
<td>c</td>
<td>d</td>
<td>d</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>b</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>d</td>
<td>d</td>
<td>d</td>
<td>c</td>
</tr>
<tr>
<td>Difficult</td>
<td>c</td>
<td>c</td>
<td>c</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>c</td>
<td>c</td>
<td>c</td>
<td>e</td>
</tr>
<tr>
<td></td>
<td>e</td>
<td>e</td>
<td>d</td>
<td>d</td>
<td>c</td>
<td>c</td>
<td>b</td>
<td>b</td>
<td>B</td>
<td></td>
</tr>
</tbody>
</table>

Note.


In the questionnaire phase, the informants completed a Likert questionnaire designed to help them reflect on why they ranked the essays in the way that they did and describe their ranking processes in a way that provides insight into the features that they felt influenced their perceptions of difficulty when reading the exemplars from the rhetorics. When examining the informants’ responses, we found that the informants believed that the two features measured by the Lexile readability formula (i.e., vocabulary and sentence length) and 12 other features
influenced their perceptions of difficulty when reading the exemplars taken from the rhetorics (Table 5).

Table 5. Informants’ Responses to the Questionnaire

<table>
<thead>
<tr>
<th>Features</th>
<th>Annie</th>
<th>Ben</th>
<th>Dan</th>
<th>Eve</th>
<th>Harold</th>
<th>Jacob</th>
<th>Kala</th>
<th>Linda</th>
<th>Marsha</th>
<th>Nelson</th>
<th>F</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Vocabulary</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>10</td>
<td></td>
<td>100%</td>
</tr>
<tr>
<td>2. Vocabulary in Context</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>9</td>
<td>90%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Logical Organization</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>8</td>
<td>80%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Structure</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>8</td>
<td>80%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Background</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>7</td>
<td>70%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Interest</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>7</td>
<td>70%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Signal Words</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>7</td>
<td>70%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Sentence Length</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>5</td>
<td>50%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Overall Length</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>5</td>
<td>50%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Paragraph Length</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>4</td>
<td>40%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Supp Materials 1 a</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>3</td>
<td>30%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Format</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>2</td>
<td>20%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Supp Materials 2 b</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>2</td>
<td>20%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Titles</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>1</td>
<td>10%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note.

a Supp Materials 1. Supplementary Materials (Advance organizers: Introductions prior to the writing text.)
b Supp Materials 2. Supplementary Materials (Pre- and post-questions: The questions before and/or after the text.)

In the second stage, the informants participated in semi-structured retrospective interviews to elaborate on the quantitative findings of the difficulty ranking and questionnaire procedure. By examining the informants’ responses, a deeper understanding emerged regarding how the informants perceived the features both measured and not measured by the Lexile readability formula. In total, they reported that they believed that 16 features influenced their perceptions of difficulty, namely, the two Lexile readability formula measures (i.e., vocabulary and sentence length) and 14 other measures (Table 6). The data are listed alphabetically.
Table 6. Features the Respondents Perceived as Influencing Text Difficulty

<table>
<thead>
<tr>
<th>Features Considered by the Lexile Readability Formula</th>
<th>Features Not Considered by the Lexile Readability Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Sentence Length</td>
<td>1. Background Knowledge</td>
</tr>
<tr>
<td>2. Vocabulary</td>
<td>2. Formatting</td>
</tr>
<tr>
<td></td>
<td>3. Grammar</td>
</tr>
<tr>
<td></td>
<td>4. Interest</td>
</tr>
<tr>
<td></td>
<td>5. Logical Organization</td>
</tr>
<tr>
<td></td>
<td>6. Overall Length</td>
</tr>
<tr>
<td></td>
<td>7. Paragraph Length</td>
</tr>
<tr>
<td></td>
<td>8. Punctuation</td>
</tr>
<tr>
<td></td>
<td>9. Signal Words</td>
</tr>
<tr>
<td></td>
<td>10. Structure</td>
</tr>
<tr>
<td></td>
<td>11. Supp Materials 1&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>12. Supp Materials 2&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>13. Titles</td>
</tr>
<tr>
<td></td>
<td>14. Vocabulary in Context</td>
</tr>
</tbody>
</table>

Note.

<sup>a</sup> Supp Materials 1. Supplementary Materials (Advance organizers: Introductions prior to the writing text.)

<sup>b</sup> Supp Materials 2. Supplementary Materials (Pre- and post-questions: The questions before and/or after the text.)

The results of the interviews further showed that when considering the difficulty of the reading exemplars taken from the rhetorics, as a group, the informants perceived each of these 16 features to be (a) sometimes a primary feature (i.e., an isolated feature), (b) sometimes a conjoined feature (i.e., consisting of two or more associated entities in which the second entity impacts the first entity), and (c) sometimes neither a primary nor a conjoined feature but simply influential.

Specifically, the results showed that the informants considered 15 of these features to be primary features (Table 7). The data are listed first by frequency and then alphabetically.

Table 7. Features the Informants Reported to Be Primary

<table>
<thead>
<tr>
<th>Feature</th>
<th>%</th>
<th>Feature</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Vocabulary</td>
<td>90%</td>
<td>9. Formatting</td>
<td>30%</td>
</tr>
<tr>
<td>2. Vocabulary in Context</td>
<td>90%</td>
<td>10. Overall Length</td>
<td>30%</td>
</tr>
<tr>
<td>3. Logical Organization</td>
<td>80%</td>
<td>11. Supp Materials 1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>30%</td>
</tr>
<tr>
<td>4. Signal Words</td>
<td>80%</td>
<td>12. Grammar</td>
<td>20%</td>
</tr>
<tr>
<td>5. Interest</td>
<td>70%</td>
<td>13. Supp Materials 2&lt;sup&gt;b&lt;/sup&gt;</td>
<td>20%</td>
</tr>
<tr>
<td>6. Background Knowledge</td>
<td>60%</td>
<td>14. Paragraph Length</td>
<td>10%</td>
</tr>
<tr>
<td>7. Structure</td>
<td>50%</td>
<td>15. Sentence Length</td>
<td>10%</td>
</tr>
<tr>
<td>8. Titles</td>
<td>40%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Note.

a Supp Materials 1. Supplementary Materials (Advance organizers: Introductions prior to the writing text.)
b Grammar was not included on the original questionnaire.
c Supp Materials 2. Supplementary Materials (Pre- and post-questions: The questions before and/or after the text.)

The results of the interviews additionally showed that one or more of the informants in the group believed that 12 features affect the readability of the exemplars in the rhetorics as conjoined features. Vocabulary, for example, was viewed as influenced by background knowledge and interest (Table 8). The data are listed first by frequency and then alphabetically.

**Table 8. Features the Informants Reported to Be Conjoined (Influenced by Other Features)**

<table>
<thead>
<tr>
<th>Conjoined Features</th>
<th>%</th>
<th>Influential Features</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Vocabulary</td>
<td>70%</td>
<td>Background Knowledge</td>
<td>60%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interest</td>
<td>20%</td>
</tr>
<tr>
<td>2. Sentence Length</td>
<td>50%</td>
<td>Vocabulary</td>
<td>30%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grammar</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Punctuation</td>
<td>10%</td>
</tr>
<tr>
<td>3. Overall Length</td>
<td>40%</td>
<td>Interest</td>
<td>20%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vocabulary</td>
<td>20%</td>
</tr>
<tr>
<td>4. Logical Organization</td>
<td>30%</td>
<td>Vocabulary</td>
<td>20%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Background Knowledge</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Titles</td>
<td>10%</td>
</tr>
<tr>
<td>5. Paragraph Length</td>
<td>30%</td>
<td>Vocabulary</td>
<td>30%</td>
</tr>
<tr>
<td>6. Interest</td>
<td>20%</td>
<td>Logical Organization</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Overall Length</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sentence Length</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vocabulary</td>
<td>10%</td>
</tr>
<tr>
<td>7. Titles</td>
<td>20%</td>
<td>Vocabulary</td>
<td>20%</td>
</tr>
<tr>
<td>8. Vocabulary in Context</td>
<td>20%</td>
<td>Logical Organization</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Paragraph Length</td>
<td>10%</td>
</tr>
<tr>
<td>9. Background Knowledge</td>
<td>10%</td>
<td>Vocabulary</td>
<td>10%</td>
</tr>
<tr>
<td>10. Grammar</td>
<td>10%</td>
<td>Sentence Length</td>
<td>10%</td>
</tr>
</tbody>
</table>

The results of the interviews further demonstrated that one of the informants considered one feature (i.e., punctuation) simply to be a feature that influenced another feature (i.e., punctuation influences the difficulty caused by long sentences).

**Discussion and Conclusions**

Because rhetorics provide apprentice writers with genre-specific models, reading genre-specific materials is an important part of the reading-writing relationship (Grabe, 2003, 2016),
and it is important to select materials that are readable for the intended users (Allington, 2002). This study provides data that can practically and theoretically inform writing center staff (when choosing texts for writing center self-access libraries), teachers of writing (when choosing texts for classroom use), members of the publishing industry (when choosing texts to market), and the research community as a whole. These data are as follows.

The data from the informants’ reports support and advance existing research. Vocabulary, for example, was reported to be a primary feature by 90% of the informants and a conjoined feature by 70%. Vocabulary was further reported to influence seven of the 16 features found to be influential in this study. Sentence length was also reported to be a primary feature, a conjoined feature, and a feature that was seen to influence three other features. Based on these data, it can be concluded that these two features (i.e., the features considered by the Lexile readability formulae) are, as other scholars have argued, strong predictors of the difficulty that students have with texts (see Dubay, 2007). However, it can also be seen that these two features are only two of the 16 features that the informants believed to contribute to the difficulty that they have when reading exemplars from rhetorics.

Overall, the results of this study support other scholars’ work that argues that although vocabulary and sentence length are strong predictors of students’ difficulty with texts, other features need to be considered when selecting texts (Baker, 2019; Chall & Dale, 1995; Fry, 2002; Gunning, 2003; Lexile, 2010; Meyer, 2003; Weaver, 2000; Zakaluk & Samuels, 1988). Therefore, this study reiterates, as other authors have argued, that readability formulae, or specifically for the purposes of this study, the Lexile readability formulae, is a good starting point in the text selection process, but other factors should be considered when deciding which books to choose (Lexile, 2010).

16. Accepting that other features beyond the features explored by the Lexile readability formulae need to be considered, this study has furthered our knowledge about what contributes to readability (regarding the features measured and not measured by the Lexile readability formula) when examining the exemplars found in rhetorics, the results of which inform a checklist that
those who wish to select rhetorics (or exemplars of paragraphs and essays therein) can use during the selection process (Table 9).

Table 9. A Checklist Which Can Be Used as Part of the Sequential Process of Selecting Rhetorics or the Exemplars Therein

<table>
<thead>
<tr>
<th>Features Considered by the Lexile Readability Formula</th>
<th>Features Not Considered by the Lexile Readability Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Sentence Length</td>
<td>1. Background Knowledge</td>
</tr>
<tr>
<td>2. Vocabulary</td>
<td>2. Formatting</td>
</tr>
<tr>
<td></td>
<td>3. Grammar</td>
</tr>
<tr>
<td></td>
<td>4. Interest</td>
</tr>
<tr>
<td></td>
<td>5. Logical Organization</td>
</tr>
<tr>
<td></td>
<td>6. Overall Length</td>
</tr>
<tr>
<td></td>
<td>7. Paragraph Length</td>
</tr>
<tr>
<td></td>
<td>8. Punctuation</td>
</tr>
<tr>
<td></td>
<td>9. Signal Words</td>
</tr>
<tr>
<td></td>
<td>10. Structure</td>
</tr>
<tr>
<td></td>
<td>11. Supp Materials 1&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>12. Supp Materials 2&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>13. Titles</td>
</tr>
<tr>
<td></td>
<td>14. Vocabulary in Context</td>
</tr>
</tbody>
</table>

<sup>a</sup> Supp Materials 1. Supplementary Materials (Advance organizers: Introductions prior to the writing text.)

<sup>b</sup> Supp Materials 2. Supplementary Materials (Pre- and post-questions: The questions before and/or after the text.)

To use this checklist, the following procedure is suggested.

Step 1: (a) Assess the readability level(s) of the texts by using the Lexile readability formula (by examining each exemplar separately, as each exemplar within a rhetoric is an individually authored, separate piece of writing), (b) examine students’ reading levels with a standardized reading assessment paired with Lexiles (e.g., the SRI) and (c) compare the two. The first two steps of Step 1 can be reversed.

Step 2: Select texts that are appropriate for students’ Lexile levels and then complete a subjective analysis of the exemplars in the potential rhetorics using the checklist. Please note that, again, each exemplar must be considered separately, as each exemplar within a rhetoric is an individual piece of writing. As you complete the checklist, consider each feature in the checklist and how each is conjoined with or influences other features.
Limitations

Consistent with the nature of qualitative theory, we purposively selected the informants to best help us understand the problem (Creswell & Creswell, 2018), and we adhered to the sample size suggested by Kvale (1996) that is commonly used in interview studies, namely, 15 ± 10. Given our sample size, however, investigations with larger samples may find different weightings regarding the features identified in this study. As such, it is hoped that this paper will lay a foundation for future investigations and provide preliminary data for further study.

Suggestions for Future Study

The information provided in this study furthers the literature in ways that can be useful to writing center staff, teachers of writing, members of the publishing industry, and the research community, but the reported data raise additional questions that merit investigation.

In the difficulty ranking procedure, we asked the informants to rank the essays from easiest to most difficult. This procedure prepared the informants for the next phase: the questionnaire. We provided the data from the difficulty ranking phase to help the reader follow the study, not as an attempt to validate or invalidate the use of the Lexile readability formula with this population or rhetorics. For such interpretations to be made, additional explorations are needed.

During the difficulty ranking and questionnaire phases and the following interview stage, we found that the informants (as a group) perceive a total of 16 features as contributing to text difficulty when reading exemplars from rhetorics, but each informant reported a different mix of features. Examining these data, one question arises: What prompted the variance in the features reported by the informants?

A broader potential focus also presents itself, and this is in the area of the generalizability of the study’s findings and the aforementioned questions. Publisher catalogues in other non-North American regions show similar available titles (i.e., rhetorics) and writing centers and the self-access libraries therein are becoming commonplace in Asia. Thus, the hypothesis can be posed that individuals who select rhetorics (or the material therein) for ELL contexts and populations could make productive use of the data provided in this study and the resulting checklist when they employ hybrid approaches to select rhetorics for use in writing center self-access libraries.

Replication studies are also encouraged in order to explore the experiences of other populations at other institutions in non-North American settings with rhetorics (as well as with other genre). This article was written in such a way to encourage such explorations, i.e., with a detailed
literature review, methodology, and results. That is because, as Baker (2019) noted, although readability investigations regarding rhetorics may have run their course in the North American context, they are only just beginning to be undertaken in the East and many other parts of the world; thus, new exploration is welcome for the texts and populations in these contexts.
References


Preprint


Ercetin, E. (2010). Effects of topic interest and prior knowledge on text recall and annotation use in reading a hypermedia text in the L2. *ReCALL, 22*(2), 228-246. doi:10.1017/S0958344010000091


Florencio, D. C. (2004). *The role of prior background knowledge in the reading comprehension of EFL Brazilian college students and American college students* (Doctoral Dissertation). Available from Proquest Dissertations and Theses Database. (UMI No. 305149503)


Fox, D. L. (1978). *The reading achievement levels of college freshman enrolled in selected course compared with the readability levels of textbooks in those assigned courses* (Doctoral Dissertation). Available from Proquest Dissertations and Theses Database. (UMI No. 7816608)


Freedle, R., & Kostin, I. (1992). *The prediction of GRE reading comprehension item difficulty for expository prose passages for each of three item types: main ideas, inferences and explicit statements*. Retrieved from ERIC database. (ED385561)


He, T. (2000). *Comprehending EFL reading comprehension: The interplay of cultural schemata and goal orientations on on-line reading strategies and reading comprehension in an EFL (Taiwan) context* (Doctoral Dissertation). Available from Proquest Dissertations and Theses Database. (UMI No. 9980991)


Morrison, L. (1978). *The relations up between the reading ability levels of freshman college students and the readability levels of required English textbooks* (Doctoral Dissertation). Available from Proquest Dissertations and Theses Database. (UMI No. 7909668)


Appendix

Questionnaire (in English)

Please read the instructions carefully before you begin. If you have any questions, please ask the proctor.

Instructions: Now that you have ordered the texts from 1 (easiest) to 5 (most difficult), please explain your reasons for arranging the texts in the way you did by completing the questionnaire below.

1. Vocabulary: The number of unfamiliar, abstract, figurative, or technical words in the text influenced my decision about how to arrange the texts in the way that I did.
   a. Strongly Agree
   b. Agree
   c. Neither Agree nor Disagree
   d. Disagree
   e. Strongly Disagree

2. Vocabulary in Context: How well the words or sentences surrounding the words I did not know helped me to understand them influenced my decision about how to arrange the texts in the way that I did.
   a. Strongly Agree
   b. Agree
   c. Neither Agree nor Disagree
   d. Disagree
   e. Strongly Disagree

3. Sentence Length: How many words were in each sentence of each text influenced my decision about how to arrange the texts in the way that I did.
   a. Strongly Agree
   b. Agree
   c. Neither Agree nor Disagree
   d. Disagree
   e. Strongly Disagree

4. Background Knowledge Required: How familiar I was with the topic of each text influenced my decision about how to arrange the texts in the way that I did.
   a. Strongly Agree
   b. Agree
   c. Neither Agree nor Disagree
   d. Disagree
   e. Strongly Disagree

5. Interest: How interested I was in the topic of each text influenced my decision about how to arrange the texts in the way that I did.
   a. Strongly Agree
   b. Agree
   c. Neither Agree nor Disagree
   d. Disagree
   e. Strongly Disagree

6. Titles: How well the title of each text described each text influenced my decision about how to arrange the texts in the way that I did.
   a. Strongly Agree
b. Agree
c. Neither Agree nor Disagree
d. Disagree
e. Strongly Disagree

7. Overall Length: How many total words were in each text influenced my decision about how to arrange the texts in the way that I did.
a. Strongly Agree
b. Agree
c. Neither Agree nor Disagree
d. Disagree
e. Strongly Disagree

8. Paragraph Length: How many words in each paragraph of each text influenced my decision about how to arrange the texts in the way that I did.
a. Strongly Agree
b. Agree
c. Neither Agree nor Disagree
d. Disagree
e. Strongly Disagree

9. Logical Rhetorical Organization: How the ideas were arranged in each text to help them flow logically from one another influenced my decision about how to arrange the texts in the way that I did.
a. Strongly Agree
b. Agree
c. Neither Agree nor Disagree
d. Disagree
e. Strongly Disagree

10. Structure: How well each text was organized (e.g., an introductory paragraph which contains a clear thesis statement, body paragraphs that contain topic sentences and supporting details, and a concluding paragraph) influenced my decision about how to arrange the texts in the way that I did.
a. Strongly Agree
b. Agree
c. Neither Agree nor Disagree
d. Disagree
e. Strongly Disagree

11. Signal words: Whether the text contained words that indicated the flow of information (e.g., first, next, finally, etc.) influenced my decision about how to arrange the texts in the way that I did.
a. Strongly Agree
b. Agree
c. Neither Agree nor Disagree
d. Disagree
e. Strongly Disagree

12. Punctuation: The way the sentences were punctuated--the use of periods (.), question marks (?), exclamation marks (!), commas (,) colons (:), semicolons (;), dashes/hyphens (-), ellipsis (…), etc.--in each text influenced my decision about how to arrange the texts in the way that I did.
a. Strongly Agree
b. Agree
c.Neither Agree nor Disagree
d.Disagree
e.Strongly Disagree

13. Format: The physical appearance of the text (i.e., font, type size, spacing—e.g., single/double, line length) influenced my decision about how to arrange the texts in the way that I did.
   a.Strongly Agree
   b.Agree
c.Neither Agree nor Disagree
d.Disagree
e.Strongly Disagree

14. Supplementary Materials 1: The introductions prior to the writing sample influenced my decision about how to arrange the texts in the way that I did.
   a.Strongly Agree
   b.Agree
c.Neither Agree nor Disagree
d.Disagree
e.Strongly Disagree

15. Supplementary Materials 2 (pre- and post-questions): The questions before and/or after the writing sample influenced my decision about how to arrange the texts in the way that I did.
   a.Strongly Agree
   b.Agree
c.Neither Agree nor Disagree
d.Disagree
e.Strongly Disagree