

## The Effects of Text Length on the Readability of Model Essays

John R. Baker  
[drjohnrbaker@tdtu.edu.vn](mailto:drjohnrbaker@tdtu.edu.vn)  
Faculty of Foreign Languages,  
Ton Duc Thang University, Vietnam

### ABSTRACT

Providing self-access libraries and the texts therein (anthologies of model essays) has long been cited as an essential part of university writing center services, as the reading of model essays has been found to positively affect the reading-writing relationship. When choosing such texts, readability is typically measured via quantitative readability formulae (e.g., the Lexile Readability Formula). However, this practice only measures two (i.e., semantic, syntactic) of the many features that influence readability, leaving others (e.g., text length) unexplored. To address this, this article reports the findings of an exploratory mixed-methods study conducted in an Asian university writing center setting which showed that the informants' ranking of ease and difficulty was significantly different than the Lexile Formula and that text length had a significant positive association with this ranking. It was further found that length was viewed as (a) a primary (i.e., an isolated feature), (b) a conjoined feature (i.e., comprising two or more associated entities where the second impacts the first: interest, vocabulary), and (c) a feature which impacts the influences of other features (interest, vocabulary, and vocabulary in context). The study also offers suggestions for writing studies professionals (teachers, writing center staff) and the publishing industry that text length be included in a hybrid (quantitative-qualitative) procedure when considering the difficulty of model essays found in anthologies.

**Keywords:** model essays; text length; anthologies; readability; Lexile

### INTRODUCTION

Anthologies of model essays have long been employed in composition studies, since the 1890s, as the reading of model essays has been found to have a positive effect on the reading-writing relationship (Saengsrichan & Chaya, 2014), providing as much as a 0.70 correlation in reading-writing related gains (Grabe & Zhang, 2016). Following this, as part of the long tradition of maintaining self-access libraries in university writing centers (North, 1984), anthologies are often placed on self-access shelves (Baker, 2020). However, simply providing model texts does not guarantee learning. The texts must be readable (Allington, 2002), a phenomenon that is often explored through a three-stage quantitative procedure: (a) an examination of a text's readability levels via a readability formula (e.g., Lexile), (b) an investigation of students' readability levels via a standardized test (e.g., the Scholastic Reading Inventory), and (c) a comparison of the two. Such a procedure is often considered a good measure of readability (Dubay, 2007), as quantitative readability formulae have been shown to effectively measure two powerful indicators of text difficulty (semantic, vocabulary; syntactic, sentence length).

Considering that quantitative readability formulae are reliable and valid predictors of readability (Dubay, 2007), there have been a limited number of historical L1 quantitative investigations of the readability of anthologies ( $N = 5$ ) (Auvenshine, 1978; Cline, 1971; Dunn, 1983; Fox, 1978; Morrison, 1978). One investigation (O'Hear et al., 1992), arguing that the

quantitative procedure was overly reductive, added a follow-up step (a cline-questionnaire procedure) to qualitatively explore one additional feature, the effect of interest on students' perceptions of text difficulty.

Accepting that employing a single quantitative step that effectively but only measures two indicators (vocabulary and sentence length) is reductive (Cunningham et al., 2018), others, both scholars and readability formulae developers (Gunning, 2003), have argued that we must include a second qualitative step that employs a subjective exploration of features that quantitative readability formulae do not effectively measure in order to provide a good fit between reader and text. In line with this, accepting that readability encompasses a wide spectrum of features, and each requires in-depth study, this investigation provides an in-depth exploration of one feature that has been much understudied (e.g., text length), in how it affects readers' perceptions of readability, both as an independent and conjoined feature.

## LITERATURE REVIEW

Historical text length literature regarding the influence length has on reading comprehension can be traced back to the late 1800s prescriptive discussions of text length (Earle, 1890; Scott & Denny, 1895) as well as Lewis' (1894) doctoral dissertation, *The History of the English Paragraph*, where Lewis surveyed text lengths dating back to old English texts. Empirical explorations have been much more recent but have generally provided mixed results with regard to how text length influences students' reading experiences.

Early work with native speakers (NS) and narrative texts (novels, short stories) has shown that longer episodic length generally facilitates comprehension and recall (Rothkopf, 1965). Black and Bower (1979) illustrated this by demonstrating that a typical story grammar contains a series of episodes and that when superordinate episodes are supported by subordinate ones (which create additional length), the latter facilitates the recall of the former. Black and Bern (1981) similarly reported that additional sentences that are causally supported are better recalled than temporally connected ones. Gerot and Unsworth (2000) further illustrated that additional length supports comprehension through fuller story development.

Explorations with expository texts excerpted from textbooks, computer manuals, and standardized tests (e.g., GRE) have provided less congruent results. Reder et al. (1986), for example, demonstrated that some elaborations (e.g., conceptual) negatively impact comprehension while others (e.g., syntactic) are facilitative. Commander and Stanwyck (1997) similarly illustrated that longer texts aid comprehension, adding that shorter ones produce an illusion of knowing (comprehension). Others found the opposite to be true. Reder and Anderson (1980), for instance, showed that while additional length (i.e., details) increases interest and credibility, it can negatively affect retention of central ideas, as much as 33% (Rothkopf & Billington, 1983). More recently, additional research has further reiterated this (Daley & Rawson, 2019), showing that extended length requires more processing time, and a higher cognitive load negatively affects comprehension.

A limited number of investigations have also been undertaken with English language learners (ELLs). Work with narratives has generally shown that comprehension improves with text length using a variety of materials: newspapers (Cha, 1995), excerpts from EFL textbooks (Jalilehvand, 2011), and short stories (Gopal & Mahmud, 2019).

As with studies in NS contexts, results in ELL settings with expository texts have been less than constant. Much of the work in this area has been undertaken with texts excerpted from

standardized exams (CAST, GRE, SAT, TOEFL). Freedle and Kostin (1993) illustrated the association between difficulty and increased length. Moon (2019) similarly reported that length was a strong indicator of difficulty, whereas Mehrpour and Riazi (2004) demonstrated that length had no significant impact. On the other hand, Bae and Lee (2018) found that length facilitates comprehension, while Bae (2017) reported mixed results, adding that reader level was a moderating variable.

Work with NSs and ELLs has also illustrated text length's confluence with other features. Extended text length, for instance, has been shown to negatively affect interest and efficacy, e.g., increasing readers' perceptions of difficulty (Shariati & Bordbar, 2011). Others have similarly demonstrated that longer texts can negatively influence comprehension due to the related increased vocabulary load. Still others, however, have shown that the additional context clues and supporting detail redundancy in longer texts can facilitate vocabulary inferencing (Ames, 1966; Shokouhi & Askari, 2010).

Overall, while a limited number of investigations have explored the effect of length on text difficulty with various genres, in-depth explorations of students' perceptions of the effect of length on the readability of model essays excerpted from anthologies are lacking. This is indeed a concern as anthologies have long been used in composition studies and included in writing center self-access libraries, and the difficulty of model essays has been cited as a needed consideration for inclusion in the 200 such texts regularly in print (Bloom, 1999). This paper addresses that gap. Specifically, it explores how length (overall and paragraph length) affects postsecondary ELLs' perceptions of readability when they read model essays excerpted from anthologies.

## METHODS

This exploratory mixed-methods study explored the effects (benefits and difficulties) length (overall and paragraph length) has on undergraduate ELLs' perceptions of difficulty when they read model essays excerpted from anthologies.

## SETTING

The study was conducted at a university in northern Taiwan that maintains a writing center that serves the university's 9,000 students, to include (a) students enrolled in requisite composition and business writing courses, (b) students from courses that include writing components, and (c) those seeking assistance with other writing needs.

## RESEARCH DESIGN

This study employed an exploratory mixed-methods adaption of Creswell and Creswell's (2018) design (Figure 1)<sup>1</sup>. Following this design, the researcher (the writing center director and course instructor for one of the sections of Sophomore Composition) conducted three steps and their relevant stages: (a) Identifying Participants and Texts (model essays), (b) An Exploration of the Association between Participants' Clines and Text Length, and (c) Exploring the Benefits and Difficulties of Length for Students who Read Model Essays Excerpted from Anthologies.

---

<sup>1</sup> This article reports the results of a larger unpublished sequential, mixed-methods study (i.e., a doctoral dissertation) that explored the effects of textual features on postsecondary ELLs' perceptions of difficulty when reading model essays excerpted from anthologies. The methodology presented in this paper was adapted from the larger study.

Step 1	Step 2	Step 3
Quantitative comparison of students' reading levels (using the Scholastic Reading Inventory) and texts' readability levels (using the Lexile Readability Formula) to identify participants and texts.	Quantitative cline procedure and exploration of correlations between participants' clines and text length.	Further qualitative exploration of the results of the cline-questionnaire procedure via semi-structured retrospective interviews.

FIGURE 1. Exploratory mixed-methods research design

**IDENTIFYING PARTICIPANTS AND TEXTS (MODEL ESSAYS)**

To identify participants and texts for this study, three stages were undertaken: (a) an examination of the target sample's reading levels, (b) an examination of the anthologies' (i.e., the model essays therein) Lexile levels, and (c) a comparison of the two.

To identify potential participants, the Scholastic Reading Inventory (SRI) (a Lexile correlated exam), was administered to a purposive sample ( $N = 91$ ) of students who regularly visited the writing center. Considering sample sizes for homogeneous purposive samples ( $15 \pm 10$ , Kvale, 1996), a cluster sample of informants ( $n = 14$ ) was identified according to their SRI scores (828–928L), which identified them as being able to compare a wide range of model essays to aid the researcher to holistically explore the research question (Merriam, 1991).

The identified participants were queried by e-mail for consent to a follow-up post-course interview. Twelve informants assented and were provided pseudonyms (females,  $n = 7$ ; mean age = 20.14 years), males ( $n = 5$ ; mean age = 20.8 years). Eleven reported to the test site, and 10 completed the procedures, thus providing useful data. The sample composition was indicated by the students' Lexile measures rather than any purposeful intent of the sampling procedure (e.g., age, ethnicity, gender) (Table 1).

TABLE 1. Participants

Pseudonym	Ethnicity	Gender	Age	Lexile Measures
Annie	Taiwanese	Female	20	928L
Ben	Taiwanese	Male	21	858L
Cara	Taiwanese	Female	20	828L
Dan	Taiwanese	Male	20	870L
Eve	Taiwanese	Female	20	861L
Harold	Taiwanese	Male	21	837L
Jacob	Taiwanese	Male	21	869L
Kala	Taiwanese	Female	20	864L
Linda	Taiwanese	Female	20	892L
Marsha	Taiwanese	Female	20	877L
Nelson	Taiwanese	Male	21	869L

To identify the model essays for the study, anthologies available on the local market ( $N = 12$ ) and the model essays therein ( $N = 893$ ) were measured with the Lexile Analyzer, an online tool that applies the Lexile Formula to texts. Next, to determine which model essays are accessible to the selected participants' reading levels, the participants' reading levels and model essays'

Lexile levels were compared. Following this, five model essays (range 610-1010L) were purposively selected to be below, within, and slightly above the informants' Lexile range (828-928L) (see Table 2). This number ( $N = 5$ ) was, in keeping with face validity, selected as it provided enough variety for the informants to engage in reflective comparisons and discussion of their essay ranking (via the cline/questionnaire procedures and semi-structured interviews), but the time required was short enough to avoid informant fatigue.

TABLE 2. Model essays chosen for the study

Model Essays	Lexile Measures	Overall Length	Number of Paragraphs	Length of Paragraphs
<i>Guide to proper hand-washing technique</i> <sup>a</sup>	610L	513	11	46
<i>Salvation</i> <sup>b</sup>	740L	887	15	46
<i>A view from the bridge</i> <sup>b</sup>	810L	1043	48	22.93
<i>Freedom and security</i> <sup>a</sup>	910L	448	10	44.8
<i>Grammy rewards</i> <sup>a</sup>	1010L	647	13	54.76

Note. <sup>a</sup>= expository, <sup>b</sup>= narrative

#### DESCRIPTION OF THE MODEL ESSAYS

The *A Guide to Proper Hand-washing Technique*, a process essay, is rated the easiest of the five essays by the Lexile Readability Formula (610L). The text contains 513 words, separated into 11 paragraphs (mean = 46 words). The *Salvation* essay, a narrative essay, rated as the second easiest (740L), contains 887 words separated into 15 paragraphs (mean = 46 words). The *A View from the Bridge* essay, rated as the third most difficult (810L), is a combination of narration and description and comprises 1,043 words separated into 48 paragraphs (mean = 23 words). *Freedom and Security*, an argumentative essay, is rated as the fourth most difficult (910L) and contains 448 words separated into 10 paragraphs (mean = 44.8 words). Finally, *Grammy Rewards*, rated as the most difficult of the five essays (1010L), a contrast essay, contains 647 words separated into 13 paragraphs (mean = 54.76 words).

#### AN EXPLORATION OF THE ASSOCIATION BETWEEN PARTICIPANTS' CLINES AND TEXT LENGTH

Following the identification of the participants and texts, associations between length and students' perceptions of ease or difficulty were explored using an untimed cline-questionnaire procedure followed by analyses of the resulting data using descriptive statistics (mean ranking) as well as inferential statistics (the Friedman Test and Pearson Correlation). In the cline phase, the participants were presented with five essays in random order. They read the texts and constructed a cline (an arrangement of the model essays from easiest to most difficult). After that, the Friedman test was utilized to explore the students' rankings, and the Pearson Correlation was employed to identify associations between the length variables (overall and paragraph length) and the participants' clines.

Following the cline procedure, a five-point Likert questionnaire (strongly agree to strongly disagree) was administered to encourage the informants to reflect on why they ranked the essays in the sequence they did and explain this in a way that would provide insight into what factors

other than those measured by the Lexile Readability Formula (i.e., length, overall and paragraph length) contributed to their perceptions of difficulty. The questionnaire addressed a variety of features related to readability, two of which explored length, i.e.,

Overall length: How many words were in each text influenced my decision to arrange the texts in the way I did, and

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.

The questionnaire was translated from English and administered in the students' L1 (i.e., Mandarin) to avoid language ambiguities. A second translator also verified the translation. It was then pre-tested with participants not included in the sample ( $N = 2$ ). Afterward, it was administered to the target sample.

Semi-structured retrospective interviews were then administered to triangulate the data from the cline-questionnaire procedure (Creswell & Creswell, 2018). The interviews continued until data saturation was reached (mean = 32.5 minutes, range 19.3-57.4 minutes) and were video and audiotaped, transcribed, and member checked. To further avoid language ambiguities, a research assistant (translator) assisted. The resulting data were then coded using Erlandson et al.'s (1993) emergent category analysis procedure, and a second-level group debate procedure was included to add further strength and fertility to the entire analysis.

## **RESULTS AND DISCUSSION**

### **RESULTS**

This exploratory mixed-methods study explored the effects (benefits and difficulties) length (overall and paragraph length) has on undergraduate ELLs' perceptions of difficulty when they read model essays excerpted from anthologies. These results are elaborated on in the following sections: (a) Results of the Cline Procedure, (b) Results of the Exploration of the Association between Participants' Clines and Text Length, and (c) Results of the Questionnaire and Qualitative Semi-Structured Retrospective Interviews.

### **RESULTS OF THE CLINE PROCEDURE**

The results of the cline procedure are shown in Table 3. The essays, labeled a-e, are shown according to the informants' rankings from easiest to difficult.

TABLE 3. The results of the informants' cline ordering

	Annie	Ben	Dan	Eye	Harold	Jacob	Kala	Linda	Marsha	Nelson
<b>Easiest</b>	a	a	a	a	a	a	a	a	a	a
	d	d	b	c	d	e	e	e	e	d
	b	b	e	e	e	d	d	d	d	c
	c	c	c	b	b	b	b	c	c	e
<b>Difficult</b>	e	e	d	d	c	c	c	b	b	b

Note. a. *A Guide to Proper Hand-washing Technique* (610L), b. *Salvation* (740L), c. *A View from the Bridge* (810L), d. *Freedom and Security* (910L), e. *Grammy Rewards* (1010L)

The descriptive results of the cline ordering (mean rank table) (see Table 4) demonstrate that the *A Guide to Proper Hand-washing Technique* essay (610L) (mean = 1.0, out of 5.0) was found to be the easiest, followed by *Freedom and Security* (910L) (mean = 3.0), *Grammy Rewards* (1010L) (mean = 3.1), *Salvation* (740L) (mean = 3.90), and *A View from the Bridge* (810L) (mean = 4.0).

TABLE 4. Descriptive statistics

Model Essay	N	Mean (SD)	Min	Max
a. <i>A Guide to Proper Hand-washing Technique</i> (610L)	10	1.00 (0.00)	1.00	1.00
d. <i>Freedom and Security</i> (910L)	10	3.00 (1.15)	2.00	5.00
e. <i>Grammy Rewards</i> (1010L)	10	3.10 (1.2)	2.00	5.00
b. <i>Salvation</i> (740L)	10	3.9 (0.99)	2.00	5.00
c. <i>A View from the Bridge</i> (810L)	10	4.00 (0.94)	2.00	5.00

The results of the Friedman Test<sup>2</sup> reveal that there is a significant difference in the ranking of each model essay ( $X^2(4) = 23.28, p < .001$ ), thus demonstrating that the informants made definitive choices in their rankings.

Examining the predictive Lexile measures alongside the student rankings (Table 5), the results illustrate that the mean student ranking (a, d, e, b, c) runs contrary to the Lexile results (a, b, c, d, e).

TABLE 5. Comparison of Lexile and participants' rankings

Lexile Cline	Students' Clines	Mean (SD)
a. <i>A Guide to Proper Hand-washing Technique</i> (610L)	a. <i>A Guide to Proper Hand-washing Technique</i> (610L)	1.00 (0.00)
b. <i>Salvation</i> (740L)	d. <i>Freedom and Security</i> (910L)	3.00 (1.15)
c. <i>A View from the Bridge</i> (810L)	e. <i>Grammy Rewards</i> (1010L)	3.10 (1.2)
d. <i>Freedom and Security</i> (910L)	b. <i>Salvation</i> (740L)	3.9 (0.99)
e. <i>Grammy Rewards</i> (1010L)	c. <i>A View from the Bridge</i> (810L)	4.00 (0.94)

<sup>2</sup> Normality of the data was not demonstrated and thus the Friedman test was employed.

The results of the Pearson Correlation<sup>3</sup> indicated no significant association between the Lexile and the informants' mean ranking ( $r = 0.20, p = 0.75$ ). A significant positive association ( $r = 0.94, p = 0.017$ ) between the informants' mean rankings and overall length, however, was found. Yet, no positive association between the informants' mean ranking and paragraph length was identified ( $r = -0.618, p = 0.267$ ).

## RESULTS OF THE QUESTIONNAIRE AND QUALITATIVE SEMI-STRUCTURED RETROSPECTIVE INTERVIEWS

### RESULTS OF THE QUESTIONNAIRE

After the completion of the clines, a closed-response questionnaire was administered. The result (frequency analysis) indicated that five of the informants found Overall Length to contribute to their perceptions of the readability of the model essays excerpted from the anthologies, whereas Paragraph Length was found to be the contributor by four of the informants (see Table 6).

TABLE 6. Informants' responses to the questionnaire

Features	Annie	Ben	Dan	Eve	Harold	Jacob	Kala	Linda	Marsha	Nelson	<i>f</i>
1. Overall Length		x		x	x			x	x		5
2. Paragraph Length	x	x			x				x		4

### RESULTS OF THE QUALITATIVE SEMI-STRUCTURED RETROSPECTIVE INTERVIEWS

After the informants completed the questionnaire, they engaged in semi-structured retrospective interviews. An analysis of the informants' qualitative responses demonstrated that the informants as a group perceived both overall and paragraph length to be a primary feature (i.e., an isolated feature) as well as a conjoined feature (i.e., consisting of two or more associated entities where the second impacts the first). These results are discussed in the following sections.

### PRIMARY FEATURES

Both overall length and paragraph length were cited as primary features.

#### OVERALL LENGTH

Overall length (i.e., the number of words in a text) was reported to be a primary feature by three of the informants (Ben, Harold and Linda) (Table 7). Two of the informants (Ben and Harold) contrasted the two narrative essays, *Salvation* (886 words) and *A View from the Bridge* (1,043 words), both citing overall length as one factor to show why they found the first more difficult than the second. These results run contrary to literature which found that greater length in narrative texts facilitates reading ease (Black & Bower, 1979).

Another informant (Linda) cited overall length when contrasting an expository essay (i.e., *Grammy Rewards*) and a narrative essay (i.e., *Salvation*), the second and fifth essays in her cline.

<sup>3</sup> Normality of the data was demonstrated and thus the Pearson correlation coefficient was employed.



She reported that she found the expository essay (648 words) shorter and thus easier than the more extended narrative essay (886 words). This report is generally in accordance with research showing that shorter texts are comprehended better than longer ones (Commander & Stanwyck, 1997; Reder et al., 1986). However, Linda’s report about the difficulty of the expository and narrative essays potentially raises the question of how other confounding factors impacted this ranking and is contrary to literature that has shown narrative essays are generally easier than expository ones (Bereiter, 1978, ctd. in Calfee & Curley, 1984).

Overall length was also cited as a factor by one informant (Ben) to contrast two expository essays, the first and fifth essays in his cline (i.e., *A Guide to Proper Hand-washing Technique* and *Grammy Rewards*), the first being shorter than the next (i.e., 513 words vs. 648 words). Ben’s report is in accordance with literature that has found that students tend to better comprehend shorter expository texts than longer ones (Daley & Rawson, 2019; Freedle & Kostin, 1993; Reder & Anderson, 1980; Rothkopf & Billington, 1983).

#### PARAGRAPH LENGTH

Paragraph length (i.e., the number of words in each paragraph) was cited as a primary feature by one informant (Annie) (Table 7). Annie offered the general complaint that she would forget the previous paragraph if the next one was too long. Annie’s report is related to Reder and Anderson’s (1980) work that explained that students have limited working memory and may have trouble recalling texts that put too many demands on that memory. Annie’s report also agrees with empirical studies that have demonstrated that students comprehend shorter texts better than longer ones (Daley & Rawson; Freedle & Kostin, 1993; 2006; Reder & Anderson, 1980; Rothkopf & Billington, 1983).

TABLE 7. Features the informants reported to be primary

Features	Annie	Ben	Dan	Eve	Harold	Jacob	Kala	Linda	Marsha	Nelson	<i>f</i>
Overall Length		x			x			x			3
Paragraph Length	x										1

#### CONJOINED FEATURES

Both overall length and paragraph length were cited as conjoined features.

#### OVERALL LENGTH

Overall length was reported to be a conjoined feature by four informants (Ben, Eve, Marsha, and Harold). In addition, two features were reported to influence the amount of difficulty caused by overall length: (a) vocabulary and (b) interest.

For vocabulary, three of the informants (Eve, Marsha, and Harold) offered comments demonstrating they felt vocabulary negatively influenced overall length, but their reasons were markedly different. Eve and Harold explained that they thought longer essays were harder than shorter ones because of the amount of unfamiliar vocabulary contained therein. Marsha similarly explained that short essays could also be problematic if they have large amounts of unfamiliar

vocabulary. Each of these reports is supported by Buck et al. (1997), who found that vocabulary was related to overall text length difficulty.

For the area of interest, two of the informants (Ben and Harold) offered comments indicating that interest was positively and negatively influenced by overall length. Ben, for example, pointed to the *A View from the Bridge* essay and explained he felt that the essay was difficult because of its long overall length but that he was willing to read it because it “touched his heart.” On the other hand, Harold offered a divergent comment, explaining that the *Salvation* essay’s overall length reduced his interest and increased his perception of difficulty about the essay. Harold’s report is loosely supported by work that has shown that a text’s appearance can influence readers’ interest in reading an essay and their perception of difficulty (Shariati & Bordbar, 2011), making them feel the essay is formal and unapproachable if they associate the look of a text with other texts they have had negative experiences with in the past (Schrivier, 1997).

#### PARAGRAPH LENGTH

Three of the informants (Ben, Harold, and Marsha) reported one feature, vocabulary, negatively influenced the amount of difficulty caused by long paragraphs. One informant, Ben, for example, complained that he found the paragraphs in the *Grammy Rewards* essay difficult because of unfamiliar vocabulary. Ben’s report is related to the findings of Buck et al. (1997), who demonstrated that vocabulary and paragraph length are related.

Vocabulary in context was also reported to be a conjoined feature influenced by paragraph length. For example, one informant, Annie, explained that the examples following an unknown phrase (*Freedom is rooted in choice*) in the *Freedom and Security* essay helped her make sense of the phrase. Annie’s report is associated with Ames’ (1966) argument that context clues derived from supporting details, which also add length, help readers make inferences about unknown words. Annie’s report is also related to Shokouhi and Askari’s (2010) study that showed that the redundancy found in additional length facilitates ELLs’ ability to make inferences.

Taken together, the results indicated that the informants rated the essays in a significantly different order of ease and difficulty than that of the Lexile Readability Formula and both overall and paragraph length had a strong association with this ranking, albeit the association of overall length was found to be significant whereas paragraph length was not. The results further showed that the informants perceived both overall and paragraph length to be (a) a primary (i.e., an isolated feature), (b) a conjoined feature (i.e., comprising two or more associated entities where the second impacts the first: interest, vocabulary), and (c) a feature which impacts the influences of other features (interest, vocabulary in context).

#### CONCLUSION

Overall, the results further L2 readability and writing studies literature with regard to the readability of anthologies (and the model essays therein) (Auvenshine, 1978; Cline, 1971; Dunn, 1983; Fox, 1978; O’Hear et al., 1978). The findings also further literature that has suggested that features beyond those measured by the Lexile Readability Formula need to be taken into account when selecting model texts (Gunning, 2003).

Following this, this article reiterates the suggestion that using readability formulae is a good first step during the text selection process, but a second step that includes a subjective consideration of features readability formulae do not measure (e.g., length, overall and paragraph

length) needs to be included in what is known as a hybrid procedure (Baker, 2020; Gunning, 2003). Text readability is first examined via quantitative readability formulae, students' reading levels are explored using a standardized test (correlated with a readability formula), and the two data are compared, providing a first step estimate. The second, through a subjective exploration of other features, helps to extend the assessment to provide texts that are a good fit for potential readers.

In addition to providing a unique contribution to L2 readability and writing studies literature, the results have a practical implication for the text selection practices of those who choose or recommend model texts for apprenticing writers as well as for the publishing industry in constructing the 200-plus anthologies published annually. However, further exploration is still needed. First, additional explorations of the many other features that make up readability have been markedly underexplored. Second, investigations into the readability of anthologies (i.e., in the North American L1 context) received minimal attention during the 1970 and 1980s, which prematurely ceased. Third, this study explored lengths across expository and narrative genres, and thus independent investigations of length respective of each mode are encouraged.

As such, this is still a very under-explored area in L2 writing studies. Accepting that this, this article purposively provides a detailed methodology and results section in the hopes of prompting future discussions that will lead to a rich body of L2 writing center literature. In line with this, drawing on Summerfield's (1988) reflection of the early years of North American writing center literature, this territory is mainly uncharted, and thus those in L2 writing contexts are encouraged to build on this work to help map our journeys.

## REFERENCES

- Allington, R. (2002). You can't learn much from books you can't read. *Educational Leadership*, 60(3), 16–19.
- Ames, W. S. (1966). The development of a classification scheme of contextual aids. *Reading Research Quarterly*, 2(1), 57–82. <https://doi.org/10.2307/747039>
- Auvenshine, A. (1978). *A study of the readability of junior and community college textbooks used in the academic areas and the reading abilities of students using the textbooks* (UMI No. 7820661). Ph.D thesis, University Microfilms International. Available from Proquest Dissertations and Theses Database.
- Bae, M. R. (2017). *The effects of text length and question type on the reading comprehension test performance of Korean EFL readers*. Ph.D thesis, Seoul National University.
- Bae, M., & Lee, B. (2018). Effects of text length and question type on test-takers' performance on fill-in-the-blank items in Korean CSAT. *English Teaching*, 73(4). <https://doi.org/10.15858/engtea.73.4.201812.149>
- Baker, J. R. (2020). *A checklist for use with the Lexile Readability Formula when choosing materials for writing center self-access libraries*. *The Asian ESP Journal*, 16(6.1), 9-68.
- Black, J. B., & Bern, H. (1981). Causal coherence and memory for events in narratives. *Journal of Verbal Learning and Verbal Behavior*, 20(3), 267–275. [https://doi.org/10.1016/S0022-5371\(81\)90417-5](https://doi.org/10.1016/S0022-5371(81)90417-5)
- Black, J. B., & Bower, G. H. (1979). Episodes as chunks in narrative memory. *Journal of Verbal Learning and Verbal Behavior*, 18(3), 309–318. [https://doi.org/10.1016/S0022-5371\(79\)90173-7](https://doi.org/10.1016/S0022-5371(79)90173-7)

- Buck, G., Tatsuoka, K., & Kostin, I. (1997). The subskills of reading: Rule-space analysis of a multiple-choice test of second language reading comprehension. *Language Learning*, 47(3), 423–466. <https://doi.org/10.1111/0023-8333.00016>
- Calfee, R. C., & Curley, R. (1984). Structure of prose in the content areas. In J. Flood (Ed.), *Understanding reading comprehension* (pp. 161–180). International Reading Association.
- Cha, K. A. (1995). The effect of text length in the diagnosis of reading comprehension. *Journal of the Applied Linguistics Association of Korea*, 8(1), 249–275. <https://www.dbpia.co.kr/Journal/articleDetail?nodeId=NODE01852933>
- Cline, T. P. (1971). *A comparison of the readability of community college textbooks with the reading ability of the students who use them* (ED050730). ERIC. <https://files.eric.ed.gov/fulltext/ED050730.pdf>
- Commander, N. E., & Stanwyck, D. J. (1997). Illusion of knowing in adult readers: Effects of reading skill and passage length. *Contemporary Educational Psychology*, 22(1), 39–52. <https://doi.org/10.1006/ceps.1997.0925>
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Thousand Oaks: Sage.
- Cunningham, J. W., Hiebert, E. H., & Mesmer, H. A. (2018). Investigating the validity of two widely used quantitative text tools. *Reading and Writing*, 31(4), 813–833. <https://doi.org/10.1007/s11145-017-9815-4>
- Daley, N., & Rawson, K. A. (2019). Elaborations in expository text impose a substantial time cost but do not enhance learning. *Educational Psychology Review*, 31(1), 197–222. <https://doi.org/10.1007/s10648-018-9451-9>
- Dalfonos, D. (2003). Grammy rewards. In T. Cooley (Ed.), *The Norton sampler: Short essays for composition* (pp. 206–208). New York: Norton & Company.
- Dubay, W. (2007). *Unlocking language: The classic readability studies*. Impact Information.
- Dunn, J. (1983). A comparative study: College freshman reading abilities and readabilities of required texts (UMI No. EP11518). Ph.D thesis, Kean University. Available from Proquest Dissertations and Theses Database.
- Earle, J. (1890). *English prose: Its elements, history, and usage*. Smith & Elder.
- Erlanson, D. A., Harris, E. L., Skipper, B. L., & Allen, S. D. (1993). *Doing naturalistic inquiry: A guide to methods*. Thousand Oaks: Sage.
- 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* (UMI No. 7816608). Ph.D thesis, East Texas State University. Available from Proquest Dissertations and Theses Database.
- Freedle, R., & Kostin, I. (1993). The prediction of TOEFL reading item difficulty: Implications for construct validity. *Language Testing*, 10(2), 133–170. <https://doi.org/10.1177/026553229301000203>
- Gerot, L., & Unsworth, L. (2000). Exploring reading processes. In L. Unsworth (Ed.), *Researching language in schools and communities: Functional linguistic perspectives* (pp. 204–221). London: Cassel.
- Gopal, R., & Mahmud, C. T. (2019). Prose reading: The influence of text-reader factors. *Studies in English Language and Education*, 6(2), 187–198. <https://doi.org/10.24815/siele.v6i2.13367>

- Grabe, W., & Zhang, C. (2016). Reading-writing relationships in first and second language academic literacy development. *Language Teaching*, 49(3), 339–355. <https://doi.org/10.1017/S0261444816000082>
- Gunning, T. G. (2003). The role of readability in today's classrooms. *Topics in Language Disorders*, 23(3), 175–189.
- Harris, S. (2001). Freedom and security. In G. Levin (Ed.), *Prose models* (pp. 389-392). Belmont, CA: Wadsworth.
- Hughes, L. Salvation (2012). In S. V. Buscemi, & C. Smith (Eds.), *75 readings plus* (pp. 10–14). New York: McGraw-Hill.
- Jalilehvand, M. (2012). The effects of text length and picture on reading comprehension of Iranian EFL students. *Asian Social Science*, 8(3), 329–337. <https://doi.org/10.5539/ass.v8n3p329>
- Kvale, S. (1996). *Interviews: An introduction to qualitative research interviewing*. Thousand Oaks: Sage.
- Lewis, H. (1894). *The history of the English paragraph* (UMI No. 301748991). Ph.D thesis, University of Chicago. Available from Proquest Dissertations and Theses Database.
- McDonald, C. P. (2003). A view from the bridge. In T. Cooley (Ed.), *The Norton sampler: Short essays for composition* (pp. 37–41). New York: Norton & Company.
- Mehrpour, S., & Riazi, A. (2004). The impact of text length on reading comprehension in English as a second language. *Asian EFL Journal*, 3(6), 1–14. <http://hdl.handle.net/1959.14/1085840>
- Merriam, S. B. (1991). *Case study research in education: A qualitative approach*. San Francisco: Jossey.
- Moon, Y. S. (2019). *Examining the effects of test characteristics on the difficulty of an EFL high-stakes reading comprehension test*. Unpublished Master's thesis, Seoul University.
- Morrison, L. (1978). *The relationship between the reading ability levels of freshman college students and the readability levels of required English textbooks* (UMI No. 7909668). Ph.D thesis, East State Texas University. Available from Proquest Dissertations and Theses Database.
- North, S. M. (1984). *The idea of a writing center*. *College English*, 46(5), 433–446. [doi:10.2307/377047](https://doi.org/10.2307/377047)
- O'Hear, M., Ramsey, R., & Baden, W. (1992). Measuring human interest in first-year college writing textbooks. *Reading Research and Instruction*, 32(1), 64–76. <https://doi.org/10.1080/19388079209558106>
- Reder, L. M., & Anderson, J. R. (1980). A comparison of texts and their summaries: Memorial consequences. *Journal of Verbal Learning and Verbal Behavior*, 19(2), 121–134. [https://doi.org/10.1016/S0022-5371\(80\)90122-X](https://doi.org/10.1016/S0022-5371(80)90122-X)
- Reder, L. M., Charney, H. M., & Morgan, K. I. (1986). The role of elaborations in learning a skill from an instructional text. *Memory & Cognition*, 14(1), 64–78. <https://doi.org/10.3758/BF03209230>
- Rothkopf, E. Z. (1965). Some theoretical and experimental approaches to problems in written instruction. In J. Krumboltz (Ed.), *Learning and the educational process* (pp. 193–221). Rand McNally.
- Rothkopf, E. Z., & Billington, M. J. (1983). Passage length and recall with test size held constant: Effects of modality, pacing, and learning set. *Journal of Verbal Learning and Verbal Behavior*, 22(6), 667–681. [https://doi.org/10.1016/S0022-5371\(83\)90395-X](https://doi.org/10.1016/S0022-5371(83)90395-X)

- Saengsrichan, N., & Chaya, W. (2014). The effectiveness of noticing strategy instruction on developing Thai EFL students' paragraph writing quality. *UMT-POLY Journal*, 11(2), 17–26. <https://so06.tci-thaijo.org/index.php/umt-poly/article/view/29466/25329>
- Schriver, K. A. (1997). *Dynamics in document design: Creating text for readers*. New York: Wiley.
- Scott, F. N., & Denny, J. (1895). *Paragraph-writing*. Allyn & Bacon: Boston.
- Shariati, M., & Bordbar, A. (2011). Interrelationship among foreign language reading anxiety, reading proficiency and text feature awareness in the university context. *Iranian Journal of Applied Language Studies*, 1(2), 179–206. <https://doi.org/10.22111/IJALS.2011.57>
- Shokouhi, H., & Askari, S. (2010). The effect of guessing vocabulary in reading authentic texts among pre-university students. *Second Language Acquisition and Teaching*, 17, 75–89. <http://slat.arizona.edu/arizona-working-papers-second-language-acquisition-teaching>
- Summerfield, J. (1988). Writing centers: A long view. *The Writing Center Journal*, 8(2), 3–9. <https://www.jstor.org/stable/43441860>
- Traig, J. (2010). A guide to proper hand-washing technique. In M. L. Conlin (Ed.), *Patterns plus: A short prose reader with argumentation* (pp. 176–178). New York, NY: Houghton Mifflin.

#### ABOUT THE AUTHOR

John R. Baker (Ph.D) has worked with writing and self-access centers and taught writing, ESOL, and literature in the U.S.A. and Asia (Korea, Taiwan, and Vietnam). His research interests include second language writing and reading, self-access and writing center administration, research methods, and how these come together in an interdisciplinary nature. <https://orcid.org/0000-0003-3379-4751>