Journal of Literacy and Technology
Volume 11, Number 4: November 2010
ISSN: 1535-0975

The Full Stop Effect: Using Readability Statistics with Young Writers

Velma J. Beaglehole
PhD Candidate
Division of Education, Arts and Social Science
School of Education
University of South Australia
St Bernard's Road
Magill, South Australia, Australia 5072
velma.beaglehole@postgrads.unisa.edu.au

Gregory C. R. Yates
Senior Lecturer
Division of Education, Arts and Social Science
School of Education
University of South Australia
St Bernard's Road
Magill, South Australia, Australia 5072
gregory.yates@unisa.edu.au

Key Words: writing, readability, educational computing, feedback

Volume 11, Number 4: November 2010
ISSN: 1535-0975

Abstract

Readability statistics can be easily generated using word processing packages or through the Internet. In this project, 103 ten-year-old students wrote for 30 minutes on the topic of rainbows, being specifically instructed to pay attention toward punctuation. Readability statistics were generated for each product. However, initially, much of the computer-generated feedback appeared meaningless. For example, high grade levels were assigned to apparently poor writing. A closer inspection of the written products revealed 64 students had omitted full stops (periods). The insertion of full stops, in appropriate places, considerably altered the incongruous feedback information. Remediation sessions were used to teach students to recognize 'faulty' statistics generated by their own products. This brief instruction then enabled them to use the computer-generated objective feedback to make their work more readable. Addressing surface features, such as the effective insertion of full stops, frees the student to concentrate on deeper aspects of the written text.

Introduction

Writing high quality, readable text does not come easily to most young children. Many elementary school teachers express frustration at the apparent poor written products emerging from their students. However, a large body of research has indeed indicated that students will benefit from clear and strong instruction in this area (Bruning, Schraw, Norby, & Ronning, 2004; Graham & Harris, 1993, 2000; Graham & Perin, 2007; Hillocks, 1986; MacArthur, Graham, & Fitzgerald, 2006; Sawyer, Graham, & Harris, 1992; Tracy, Reid, & Graham, 2009; Troia & Graham, 2002). Teachers, on the other hand, are hard put to carry out such instruction, which may entail considerable individual tuition and feedback.

Teachers are familiar with the results of asking students to edit their work in whole class contexts. For the most part, students insist they have reread their work, yet make few, if any, changes. Writers find it difficult to locate their own errors, as they tend to read what they think is written, rather than what actually is written (Daneman & Stainton, 1993; Pilotti & Chodorow, 2009; Rogers & Graham, 2008). Reading aloud to oneself or another person has been found to be helpful (Graves, 1983; Murray, 1984). But, given the context and reality of classroom instruction, a fundamental issue concerns the relative absence of corrective feedback given to the individual. Can feedback from a computer help?

The present project was concerned with using a readily available technology as feedback to encourage young writers. The lead author has served as a key Information, Communication and Technologies (ICT) advisor in South Australian schools for 10 years, specifically encouraging teachers to develop professional skills and use the full range of ICT and Internet resources available to them and their students. The recent experience of working with Robin, a dedicated and hard working elementary school teacher, is typical of a situation seen to occur repeatedly. That is, a teacher may present with an initial openness to attempt a new skill, and a willingness to introduce a new procedure into classroom practice, but then becomes quickly discouraged at the first obstacle.

Following advice to use readability statistics, Robin, became increasingly frustrated after 'discovering' apparent grade levels, on the Flesch-Kincaid scale, ranging from 1.7 to an absurd 89.1, within his class. In addition, some of the poorer writers had higher grade levels, on this scale, than better writers. Furthermore, reading ease scores ranged from zero to 99.9. Robin could not interpret this feedback and declared the practice meaningless.

Despite readability statistics being readily available, per modern word processing packages and through a number of websites (see Appendix A), their use is restricted if the feedback is not understood. This article examines ways in which teachers, such as Robin, can learn how to use such tools effectively. Indeed, the lead author's experience is that students can be empowered to use computer-generated feedback to improve the quality of their written texts.

Volume 11, Number 4: November 2010
ISSN: 1535-0975

Readability

What is Readability?

Readability refers to the ease with which text can be read. For example, Dr Seuss books are easily read. Professional medical journals are very difficult to read. Readability is distinct from legibility, which relates to handwriting or fonts used in the presentation of text on particular backgrounds. Both the content of the written text and the use of language conventions, such as punctuation marks and paragraphing, are important in readability. A text needs to be readable for the writer to communicate the intended message. A text that is difficult to read is difficult to comprehend. In essence, readability is a prerequisite for comprehensibility.

Educators in the 1920's discovered the level of difficulty of a text could be predicted using sentence length and vocabulary (Lively & Pressey, 1923; Thorndike, 1916; Vogel & Washburne, 1928). Research in this area flourished in the 1950's and included the development of many formulas for calculating readability (Dale & Chall, 1948; Farr, Jenkins, & Paterson, 1951; Flesch, 1950; Fry, 1968; Powers, Sumner, & Kearl, 1958; Spache, 1953). In 2010, the ERIC database had over 2,700 citations under the search term *readability*.

Readability formulas have been used as an educational tool to match readers to texts in schools, and also in business and the military. In the 1950s, Associated Press hired Rudolf Flesch, creator of the Flesch Reading Ease score (FRE), to assist in making newspaper articles more readable, an issue that is apparently still alive today (Dalecki, Lasorsa, & Lewis, 2009).

However, "readability formulas are not writeability formulas" (Fry, 1989, p. 293). Readability formulas were not designed to guide writers. Klare (1984) insisted they have a prediction function, rather than a production function. They measure the readability of a piece of writing for a hypothetical (normative and virtual) audience, rather than providing instruction in text generation. Nevertheless, a

young writer might be able to profit from readability feedback, as generated easily through current technologies.

Readability Formulas

Most readability formulas are based on two measures: syntactic difficulty, and semantic difficulty. Syntactic difficulty (grammatical complexity) is measured by sentence length. Semantic difficulty (meaning or word meaning) is measured either by word length or word frequency. The number of syllables, or the number of letters, per word, calculates word length. Word frequency is based on the number of times particular words appear or by comparing the words written with a list of familiar words (Fry, 2002).

In the period prior to 1984, as researchers experimented with variables affecting readability, some 200 readability formulas were proposed (DuBay, 2004; Fry, 2002; Klare, 1984). In essence, there is no one single or 'correct' readability formula. Table 1 shows some commonly used formulas.

Flesch Readability Ease and Flesch-Kincaid Grade Level

Calculation of Flesch Reading Ease (FRE) score uses the average number of words per sentence and the average number of syllables per word. This can be expressed as FRE = 206.835 – 1.015 (total words/total sentences) – 84.6 (total syllables/total words). To calculate the FRE manually, select a 100-word written sample. Calculate the average sentence length and multiply by 1.015. Next, count the number of syllables in the 100-word sample and multiply by 0.846. Add together the average sentence length calculated, and the number of syllables calculated. Subtract this sum from 206.835. The result is the reading ease score (Flesch, 1951). Understanding how FRE scores are calculated help classroom teachers realise the impact sentence length and multisyllabic words have on the resulting score. Busy classroom teachers do not have the time to perform these laborious calculations for each child in their class. However, a computer can calculate an FRE score in a matter of seconds, making this data readily available to the classroom teacher and each student.

Volume 11, Number 4: November 2010
ISSN: 1535-0975

Table 1
Selection of Readability Formulas Depending on Words and Sentences

Readability Formula	No. words	No. sentences	No. syllables
Flesch Readability Ease	✓	✓	✓
Flesch-Kincaid Grade Level	✓	✓	✓
SMOG (McLaughlin, 1969)			✓ 3 or more syllable words in 30 sentences
FORECAST (Sticht, 1975)			✓ single syllable words per 150 words
Fry's Graph (Fry, 1968)		✓ in 100 words (3 samples)	Av syllables per 100 words (3 samples)
Gunning fog index (Gunning, 1968)		Av sentence length per 100 words	✓ no. syllables in 100 words

Note: the ticks indicate the criteria used within the formula

FRE scores can be converted into more easily understood US grades. Called the Flesch-Kincaid Grade Level (F-K GL), it is calculated using the average sentence length (ASL) and the average number of syllables per word (ASW). Counting the number of words in the text and dividing by the number of sentences calculates the ASL. Counting the total number of syllables in the text and dividing this number by the number of words calculates the ASW. The calculations are used in the following formula to generate the grade level. F-K GL = (.39 x ASL) + (11.8 x ASW) – 15.59 (DuBay, 2004). The grade levels indicate the number of years schooling required to read the text, and approximate with those used in Australia, with students turning 18 years in Year 12. On average, students in Year 4 are nine years old, turning 10, whilst those in Year 5 are 10 years old, turning eleven.

The highest readability score is around the 120 mark, where every sentence consists of one or two syllable words. *The Cat in the Hat* (Dr Seuss, 1957), for example, has a FRE of 118.1 and a Flesch-Kincaid grade level of less than one. Of the 1,603 words, only five words had more than two syllables.

The word *another* is used three times and the hyphenated words, *up-up-up* and *fun-in-a-box*, make up the other two. On average, sentences in *The Cat in the Hat* have seven words and these words, on average, have one syllable. The shorter the sentence, and the fewer syllables per word, the easier the text is to read. Flesch (1951) described a score of 90-100 as very easy, 80-90 as easy, 70-80 as fairly easy, 60-70 as standard, 50-60 as fairly difficult, 30-50 as difficult, and 0-30 as very difficult.

Applying Readability Formulas

Robin, teaching the Year 5 class, should be concerned about the writing of students with low FRE scores. A score of zero is problematic whereas a score of 99.9 indicates the writing is very easy to read. Writing scoring low grade levels are not a problem. Dr Seuss, the pen name for Theodore Seuss Geisel, was an accomplished American writer with over 60 published books. *The Cat in the Hat*, a best seller, has a remarkably low grade level. However, grade levels of 89.1 and 14.4 should alert Robin to problems in writing in his Year 5 class.

A 4,300 word article, *New Drugs, Old Drugs* published in the *Medical Journal of Australia* had a FRE of 30 (very difficult) and an F-K grade level of 13.7 (university undergraduate level) (Verma, 2010). The writer used over 1,200 complex words (28% of total word count) including words in common use such as *therapeutic, administration* and *elective*; and less common words such as *thromboprophylactic, pharmacokinetics* and *bioavailability*. The sentences were evidently lengthy, with an average of 19 words per sentence.

Coincidentally, both *The Cat in the Hat* and *New Drugs, Old Drugs* have the same number of sentences. Apart from that, their readability statistics are vastly different as Table 2 shows. Sentence length and syllables per word have a large impact on readability and the scores generated. The readability of this current article is also shown on Table 2.

Table 2
A Comparison of a Simple Text, a Complex Text and this Article

	The Cat in the Hat	*New Drugs, Old*	*The Full Stop Effect*

	(Dr Seuss, 1957)	*Drugs* (Verma, 2010)	(This article)
Flesch Reading Ease Score	118.1	30	69.7
Flesch-Kincaid Grade Level	-1.4	13.7	6.9
Word count	1603	4,319	8161
Complex word count	5	1212	1233
Sentence count	231	231	604
Average syllables per word	0.97	1.87	1.46
Average words per sentence	7	19	13.51

Note: data generated by Edit Central

Criticism of Readability Formulas

Readability formulas are not without their critics. Crossley et al. (2008), writing from a second language perspective, criticize readability formulas because they focus on the surface-level linguistic features of the text rather than the processes a reader brings to the text. Flesch (1951) acknowledged that readability was not dependent on ease of reading alone. Human interest level was also important. This interest is not in the subject but in the presentation of the subject. Consequently, Flesch provided a way of calculating the human interest level based upon the number of personal words and personal sentences in the writing, but this aspect is not represented in any of the readily available resources today.

Critics note that pronunciation of words such as *medicine*, *comparable* and *laboratory* have different syllable counts according to regional dialects. Whilst this is the case, agreeing to abide by the rule that each syllable must have a vowel sound and focusing on the written language rather than spoken language, variations are avoided, e.g. *med-i-cine, com-par-a-ble, lab-or-a-tor-y*.

It is apparent that text can be rated as 'readable' although it may not be immediately comprehensible. Incorrect spelling also makes writing difficult to read. For example:

> At home I have a poorarn and her name is coco. She likes to Play waser me and my bunes soweball and beurcs. One day my dog coco lekerd my buned sowe ball. *(Year 4 boy)*

This story was given a very easy to read FRE score of 93.1 despite being difficult to read. The story was about a Pomeranian dog, Coco, who likes to play with the writer and his bunnies, Snowball and

Beurcs. Readability formulas can also be criticized for grammatical insensitivity. A scrambled sentence is given the same readability score as an unscrambled one. Hence, by valuing surface features, such formulas can give apparently high readings on text that is 'readable' but not 'comprehensible'.

Such criticisms emphasize that readability is not the only important factor to consider when writing. Readability statistics are useful for checking the readability of writing but many additional criteria are necessary to guarantee genuinely comprehensible text.

Notwithstanding such issues, in general, readability formulas do predict how easy a particular piece of text will be for an average reader, and these are frequently based on two factors, sentence length and word length. As such, readability formulas have a heuristic value.

Generating Readability Statistics

The extent to which young students can use a computer to generate and benefit from readability statistics is unknown. The written text can be copied and pasted into Internet sites such as *Edit Central* (2009). (See Appendix A for a range of online readability calculators). In a matter of milliseconds, statistical data is presented in both graphical and text formats. Word processing packages also include readability statistics as part of their spelling and grammar checks. For example, *Microsoft Word 2007* generates readability statistics through the *Word Options* found under the *Office Button.* Selecting the *Review* ribbon and *Spelling and Grammar* in the *Proofing* section allows the writer to click through the dialogue box until the readability statistics are presented. *Microsoft Word 2007* presents counts (words, characters, paragraphs and sentences), averages (sentences per paragraph, words per sentence and characters per word) and readability (passive sentences, Flesch Reading Ease and Flesch-Kincaid Grade Level). It is not possible to copy and paste the generated statistics next to the written text unless a screen

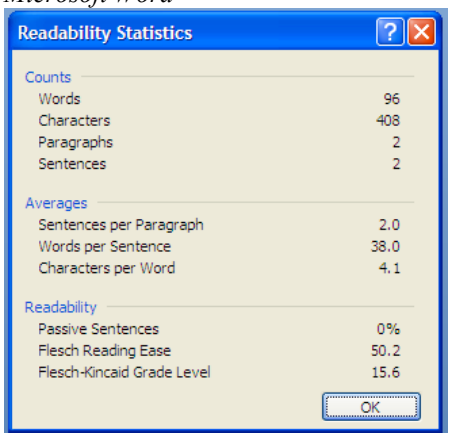
capturing process is used. This makes *Microsoft Word 2007* readability statistics less user-friendly than available internet sites.

The statistics generated are not identical due to the application of different 'rules.' For example, in Table 3, *Edit Central* counts three sentences by presuming there would be a full stop at the end of the writing. On the other hand, Microsoft *Word* counts two sentences by recognizing the full stops. Calculation of FRE scores and F-K grade levels is dependent on the number of sentences so consequently they vary too. These variations highlight the predictive nature of readability statistics. Despite appearances, the numbers generated by the formulas are not precise. Instead, they are indicative of general readability. The FRE scores and the F-K grade levels, generated by the same formulas but applying different 'rules' for sentence recognition, all indicate the text in Table 3 needs improvement in order to be more readable.

Potential for Using Readability Statistics as Feedback

It has been widely documented that conferencing a student during the writing process provides valuable feedback and improves the quality of the written product (Baker, Gersten, & Graham, 2003; Graves, 1983). A study with young adults in a business and technical writing course found readability statistics, along with qualitative feedback, helped students write more readable text (Schwartz, 1980). However, Schwartz recommended students who have errors in punctuation and sentence structure should avoid using the formulas since inflated grade levels are likely to be generated. However, an alternative view would be to help students themselves see that inflated grade level readings are created by punctuation errors that can be remediated relatively easily. In short, discrepant readability statistics can themselves be the signal that the writing product needs revision.

Table 3
Writing Sample and Readability Statistics

Year 5 Original Text	Readability Statistics generated by *Edit Central*	Readability Statistics generated by *Microsoft Word*
One bright sunny after noon my family went for a trip to the outback well my family loves to travel around Australia it isn't like just hoping in the car and driving around the coast of Australia its basically you stay at one thousand places and you visit a lot of people and places. When we got home from the out back we went to Sydney and saw the Sydney harbour bridge, the Sydney opera house. The day after that it was pouring with rain when the rain had stoped there was a beautiful coloured rainbow	Flesch reading ease score: 67.7 Automated readability index: 14.4 Flesch-Kincaid grade level: 11.8 Coleman-Liau index: 8 Gunning fog index: 15.7 SMOG index: 11.4 509 characters 408 non-space characters 404 letters/numbers 96 words 7 complex words 121 syllables 3 sentences 4.21 chars per word 1.26 syllables per word 32 words per sentence	

Writers need feedback during the process of writing as well as for the finished product. According to Hattie and Timperley (2007), effective feedback answers three major questions: (a) what are the current goals, (b) what progress is being made toward the goals, and (c) what activities need to be undertaken to make better progress? The writer aims to communicate his or her message to a reader. The writing needs to be readable. Feedback, given to the writer through an agent such as a peer, an adult, a self-directed checklist, or a computer, can enhance the

Volume 11, Number 4: November 2010
ISSN: 1535-0975

learning of the student. Readability statistics, generated easily by a computer, provide the student with task-relevant feedback about the readability of their work so far. This is powerful feedback in that it shows students the progress they are making towards producing readable writing. A low FRE score signals punctuation problems that the student can address, either by himself or herself, or by seeking help from another agent (peer, teacher, or parent).

Feedback is goal-related. Goal setting affects motivation for engaging in tasks. Adaptive goals are **s**pecific, **m**easurable, **a**chievable, **r**ealistic, and achieved in a **t**imely manner (SMART) (O'Neill, 2000). Specific goals lead to specific feedback (Locke & Latham, 1984). Students could set goals for their writing such as, "I want to polish my writing so that the FRE score is between 90 and 100 so that my buddy in the Year 2 classroom can read it. This means I need to achieve a grade level score of two or less." Once the specific goal is set, the writer can make changes to his or her written text in order to meet this SMART goal.

An elementary teacher has up to 30 students to conference in a normal classroom setting. Elementary students frequently make errors in punctuation. Teaching students how to use readability statistics to check their writing means they would be less dependent on the teacher for part of the lesson time.

The aim of the present project was to investigate ways in which teachers might be able to use readability statistics to help students edit their work. Instead of casting the incongruous statistics aside as 'nonsense,' we may consider what problems such inharmonious information might be signalling. This approach could provide support for the teacher who has the responsibility to help up to 30 students write to the best of their ability.

Research Bases Examples

Volume 11, Number 4: November 2010
ISSN: 1535-0975

Participants

One hundred and three students from a large, government school located in a middle socio-economic status district in South Australia participated in a study of children writing using computers. The participants were in their normal classroom groups and consisted of a Year 4 class, two composite Year 4/5 classes and two Year 5 classes. The median chronological age was 9 years 10 months and the mean reading age was also 9 years 10 months according to the locally used test (Waddington, 2000).

Method

In the context of a normal writing lesson, the students were asked to write original stories for 30-minutes on the theme of 'rainbows'. In the lesson, talking and collaboration was not permitted. After 25-minutes, specific instructions were given to the students requiring them to reread what they had written and add any missing punctuation, such as full stops and commas. Students written products were saved on a server.

An initial set of readability statistics was generated using *Edit Central.* Students were percentile ranked for both their FRE score and F-K grade level. A particular problem was noted running throughout the stories and across all five classes. Full stops (i.e. periods) were omitted by 64 of the 103 students, a problem evident across all five classes as indicated by a non-significant chi-square coefficient. The stories were reread by the author and full stops added at appropriate junctures. A new set of readability statistics was generated. Students were re-ranked for the new FRE score and new F-K grade level.

Results and Discussion

Initial Scores

The initial Flesch Reading Ease scores ranged from -116.1 (extremely difficult to read) to 115.5 (very easy to read). The FRE scores of the five classes did not differ significantly, $F < 1$. The FRE scores of the Year 4 and Year 5 students were not significantly different, F (1, 102) = 1.26, p = .26. The mean FRE score, uncorrected for full stops, for the 103 students was 83.65. On face value, this seems to be a reasonable readability score, falling into the easy range. However, the mean hides a wide range of scores (-116.1 to 115.5).

The initial Flesch-Kincaid grade levels ranged from -.4 (pre-school) to an impossible 84.5. The mean F-K grade level was 7.8, which seems high for students who have not yet reached that grade level. The mean F-K grade level of the Year 4 students and the Year 5 students was significantly different, F (1, 101) = 5.04, p = .03. The writing of the Year 4 students had a mean F-K grade level of 4.8 (SD = 4.0) whilst the Year 5s was 9.7 (SD = 13.56). A very capable group of Year 4 writers could possibly achieve a grade level of 4.8 but it is unlikely that a Year 5 group would be writing narratives that could only be read by students in much higher grades. Obviously, the effect of outliers actively helped to produce overall faulty readings.

Table 4 shows the readability statistics for the bottom 10 per cent of students, ranked according to their FRE scores. Aside from their FRE scores being low and their F-K grade levels being high it is obvious that the number of sentences written is low and the number of words per sentence surpasses that of a complex medical journal article (average of 19 in Verma's text). For Year 4 and Year 5 students these are incongruous readability statistics.

Table 4
Original (Incongruous) Readability Statistics as Shown by Ten Students

Student	1	2	3	4	5	6	7	8	9	10
FRE	-116.0	-51.0	8.3	43.2	45.1	57.7	58.4	59.5	62.0	62.6
F-K grade level	84.4	59.4	40.2	24.7	24.4	19.4	15.1	19.6	14.3	16.3
Sentence count	1	1	1	3	2	2	1	5	7	1

Average words per sentence	221	157	113	70	70	57	40	59	39	40
Average syllables per word	1.17	1.17	0.99	1.09	1.07	1.08	1.28	1.03	1.25	1.23
Word count	221	157	113	211	140	114	40	295	272	40

Generating New Scores by Adding Full Stops

As noted earlier, inspection of the written text revealed the curious observation that 64 out of 103 (65%) students had omitted full stops (i.e. periods). This occurred despite specific instructions given at the 25-minute mark.

Since readability statistics are driven by sentence length, it was decided to investigate the impact of adding appropriate full stops into the students' stories. Each story was reread and 460 full stops were added to the 64 stories in order to make the texts comprehensible (average of seven full stops per story, SD = 5.06). Once full stops were added, the new set of readability statistics, in some cases, showed dramatic changes in FRE scores, F-K grade levels and sentence counts. Word count and average syllables per word remain unaltered (see Table 5).

Table 5
Corrected Readability Statistics for the Same Ten Students

Student	1	2	3	4	5	6	7	8	9	10
FRE	80.0	99.8	114.2	99.3	98.4	106.7	85.4	100.7	81.7	82.9
F-K grade level	9.0	1.4	-0.4	3.2	3.9	0.6	4.7	3.8	6.7	6.7
Sentence count	8	19	13	14	8	13	3	16	14	2
Average words per sentence	28	8	9	15	18	9	13	18	19	20
Average syllables per word	1.17	1.17	0.99	1.09	1.07	1.08	1.28	1.03	1.25	1.23
Word count	221	157	113	211	140	114	40	295	272	40

Focusing on Low Incongruous Scores

A closer look at students with initial exceptionally low FRE scores revealed three students with the exceptionally low scores of -116, -51 and 8.3. The next lowest score was 43.2 but from thereon the intervals between scores were much smaller (maximum jump of 12). Only 15 per cent of students obtained an uncorrected FRE score of less than 70.

The original text of the student obtaining an FRE of -116 consisted of one sentence of 221 words with many joining words such as *and, but* and *so*. Adding full stops in appropriate places realigned his FRE score to 80. Removing some conjunctions would further enhance this score. His work was still in the first percentile, despite the addition of full stops.

The student with an original FRE score of -51 appeared to write just one sentence, but the addition of 18 full stops dramatically enhanced the readability of his work and moved him from the second to the sixty-fourth percentile. His F-K grade level dropped from an incongruous 59.4 to a realistic 1.4.

The addition of full stops to the text written by the student with an original FRE score of 8.3 showed an even more dramatic "improvement". His apparent one sentence became thirteen, and he moved from the third to the ninety-ninth percentile. His F-K grade level dropped from 40.2 to less than one. The student with an original FRE score of 43 achieved a similar result. Her apparent three sentences became fourteen, and she moved from the fourth to the sixty-third percentile. Her F-K grade level also dropped from 24.7 to 3.2. A student with an FRE score of 45.1 made a lesser percentile jump but she did move from the fifth to the 53rd percentile and the readability of her text became very easy (FRE 98.4). Her F-K grade level also dropped from 24.4 to 3.9.

In short, adding full stops into the student products had dramatic effects on readability in the case of several key students.

Comparison of FRE Scores in Uncorrected and Corrected Text

Turning attention to FRE scores by comparing the uncorrected and corrected text may help determine the FRE score a 10-year-old student should aim for in their writing. As mentioned, the FRE mean, uncorrected for full stops, hides an enormous range of scores (-116.0 to 115.5). Once corrected for full stops the mean FRE score moves up into the very easy range (96.17) and the difference between the minimum readability (69.1) and maximum readability (115.1) is much narrower (see Table 6).

Table 6
Mean Flesh Reading Ease Score for Uncorrected and Corrected Text

	Mean Flesch Reading Ease Score	S.D.	Range	Minimum	Maximum
Uncorrected for full stops	83.65	28.87	231.6	-116.0	115.5
Corrected for full stops	96.17	8.86	46.4	69.1	115.5

Even so, a readability score of 69.1 is a matter for concern as these students have a median age of 9 years 10 months and have not reached the stage of writing technical works. An investigation of the story receiving this low score showed that the student had used hyphenated words to increase the complexity of his story. They included *jolly-o, poor-o Scotto, jolly-o-land, jolly-o-boy* and *quite-o-mischievous*. These were in keeping with his story about a leprechaun on Saint Patrick's Day. This story was not one of the ones needing correction for full stops and the FRE score more accurately reflects the complexity of the story.

The student receiving the next lowest FRE score of 75.9 was also one that had not needed correction for full stops. This student wrote a 119-word story about the spirit world in which the main character was Kawakawato, a five-syllable word. Replacing this five-syllable word with a two-syllable word, Kaka, the FRE score increases to 90.8. A further increase in readability (FRE 95.8) occurs if a one-syllable word, Kale, replaces the five-syllable word. This example highlights the dependence of readability formulas on syllable counts. A teacher looking at this writing sample could reassure the

student that their work was quite readable, despite a score of 75.9, and praise the student for using a complex 5-syllable word.

Students were sorted according to their uncorrected FRE score and then again according to their corrected FRE scores. Each student received a percentile score for both the uncorrected text and the corrected text (see Table 7). For the uncorrected FRE scores the lowest 25% of students scored less than 74.4. Only the top 25% of students fell into the very easy readability range. Correcting the FRE scores raised the lowest 25% of students to scores of less than 90.4. In the corrected text, 75% of students fell into the very easy readability range.

Table 7
FRE Maximum Scores According to Percentiles

	Uncorrected FRE Score	Corrected FRE Score
25th percentile	74.4	90.4
50th percentile	89.3	96.6
75th percentile	98.3	102.3

These examples indicate two important ways to use readability statistics. Firstly, a low sentence count may indicate the omission of full stops. Secondly, if the FRE score is less than 90, for a 10-year old writer, then look at the text more closely. The FRE score may indicate that the writer has inserted very few full stops. However, it may also indicate the writer has used words with many syllables. The writer needs to check if full stops are needed. Once these are appropriately inserted, generate a new readability score. If the FRE score is still less than 90, check for multisyllabic words. *Edit Central* underlines words with three or more syllables and labels them 'complex' words, simplifying this task.

Comparison of F-K Grade Levels in Uncorrected and Corrected Text

The mean F-K Grade Level, uncorrected for full stops, for the 103 students was 7.7 (SD 11.02). As these students were either in Year 4 or in Year 5 this figure is unrealistically high. Once corrected for full stops a more realistic grade level of Grade 3 (SD 1.97) resulted (see Table 8).

Table 8
Mean Flesch-Kincaid Grade Level for Corrected and Uncorrected Text

	Mean F-K Grade Level	S.D.	Range	Minimum	Maximum
Uncorrected for full stops	7.71	11.02	85.1	-.7	84.4
Corrected for full stops	3.00	1.97	9.7	-.7	9

Sorting the students according to their uncorrected F-K grade level and then again according to their corrected F-K grade level enabled generation of percentile scores (see Table 9).

Table 9
F-K Grade Level Maximums According to Percentiles

	Uncorrected F-K Grade Level	Corrected F-K Grade Level
25th percentile	2.5	1.4
50th percentile	4.7	2.8
75th percentile	8.2	4.1

It is unrealistic to expect 25% of the Year 4 and Year 5 students would achieve grade level scores of above Grade 8. The omission of full stops had inflated the grade level dramatically. The addition of full stops lowered this to Grade 4. It is reasonable to expect 25% of Year 4 and Year 5 students could achieve a grade level above Grade 4.

Thirty-eight percent of students had no change in their Flesch-Kincaid Grade Level but 26 percent had a change (i.e. reduction) of five grades or more (see Table 10).

Table 10
Percentage of Students with Changes in F-K Grade Levels After the Addition of Full Stops

Change in Grade Level	Percentage
none	38
< 1 grade	9
1 to 2	9
2 to 3	12
3 to 4	9
5 to 10	15
> 10	11

Note: The percentage tally is more than 100 due to the rounding up effect.

It should be noted that a very low grade level is less problematic than a high grade level as students can write text for younger students still learning to read. Not unlike Dr Seuss, several students achieved a grade level of less than one but the quality of their story was high.

Correcting writing by adding full stops, in appropriate places, improves readability (a higher FRE score) and the grade level becomes more realistic. Removing some extraneous conjunctions and making shorter sentences could improve writing samples that still had a low FRE score and a high F-K grade level.

Clinical Follow-up: Teaching Students to Use *Edit Central*

The lead author returned to one of the classes to test the notion that students themselves are able to make sense out of readability feedback data. The class was shown the readability statistics generated for their work, and told there was a problem, in that the 'readability machine' could produce some numbers that did not make sense. For example, a grade level of 12 for a Year 5 writer, and 203 words but only five sentences giving an average of 40 words per sentence. It was explained how this problem could be easily fixed by adding full stops. The students were then directed to fix any similar problems, in their own work, by copying their writing into *Edit Central,* reading the feedback, and duly adding full stops.

As *Edit Central* produces 16 outcomes, students were told to focus only on three: (a) Flesch Reading Ease score (aim for 90 or more but more than 70 may be acceptable), (b) Flesch-Kincaid Grade Level (aim for your grade or less) and (c) words per sentence (aim for a lower number than was currently showing). The first outcome, Flesch Reading Ease, is at the top of the *Edit Central* list and the words per sentence are at the bottom, simplifying access to the information for students. It was apparent that every single student in the class was able to use this information and so obtain appropriate readability feedback for their own work.

However, within this session, it was noted that two students left a space before the full stop, e.g. 'He chased them. Then they ran …" It is not readily apparent to the human eye that the space is after 'them' and there is no space between the full stop and 'Then'. The computer is sensitive to these spacing issues and reflects this error in the generated scores. In one case, a Year 5 girl had left spaces before 13 of her newly inserted full stops. Thus, initially she had lowered her Flesch-Kincaid grade level from 10 to eight. However, through shifting the full stop to the appropriate position she reduced the grade level from this unrealistic eight to a more realistic two.

Extension of Use

It is clear 10-year-old students can effectively use the feedback provided by readability statistics to improve the quality of their writing. Schwartz (1980) found a combination of qualitative feedback and feedback from readability statistics enabled young adults to write text that was more readable. It seems reasonable to suggest that, with instruction, students from 10 years of age to adulthood can profit from feedback generated by readability statistics. Younger students, who are able to write complete stories may also benefit but this needs to be tested.

Current feedback from sites like *Edit Central* could be confusing for younger students who would need to scan a 16-point-long list to select the relevant data. Today's students have grown up with video games. They enjoy the challenge of working through various levels to achieve a successful outcome. There is potential for a web developer to construct a user-friendly readability statistics generator, incorporating levels, to give feedback to writers. The simplest level could provide feedback on the total number of words, the total number of sentences, the average number of words per sentence, the average number of syllables per word and the F-K grade level. Reducing the current list of 16 down to five would still provide young writers with a sense of audience, e.g. I am writing a story that could be ready by a student in grade one. Higher levels could introduce the FRE score and number of complex words.

Edit Central currently generates a colourful bar graph. A web developer could incorporate more graphics, making the site more appealing to young students. Creating the facility for students to log-on, and save their work, opens up more possibilities. Graphs showing word counts over time could be added, such as those for daily steps walked in a health web-site *10,000 steps* (Queensland Health, 2003-2010). *Edit Central* currently provides the formulae, in the Java programming language, used to generate the various readability scores. An enterprising web developer could use these to start developing a user-friendly site for young writers.

Conclusion

ICT can provide vital feedback to young writers and their teachers. Robin was one of the five teachers whose Year 5 class supplied data for this project. He now understands that grade levels of 89.1 or 14.4 indicate problems in the written text whereas grade levels of 2.8 or 1.7 are realistic for Year 5 writers. Robin knows a high FRE score is preferable to a low score. He can direct his Year 5 students to aim for an FRE score of 90, or more, whilst assuring them a score of 70 or above is acceptable if they have used multi-syllabic words. Robin expects his students to know their written text will have an F-K grade level equivalent to their current grade level or less.

Robin is confident in his use of readability statistics as they provide useful feedback to both himself and his students about their writing. He values readability statistics as offering task-specific feedback, quantifying students' work, and helping them to look at their products in an interestingly different way. His students benefit from an objective computerised report containing the number of sentences they have written, their FRE score and F-K grade level, before they meet for a conference with himself, a parent or a peer. If the sentence count is low and the FRE score, for 10-year olds, is less than 90, students know it is highly likely they need to insert more full stops in appropriate places. Students can edit their work for full stops, and remove excessive conjunctions, then recheck their readability statistics before meeting for a conference.

As a teacher, Robin knows feedback is most powerful when it follows effective instruction (Hattie & Timperley, 2007). Some students cannot benefit from the readability statistics because they first need help in constructing and recognising sentences. Such students require clear instruction rather than additional feedback. Readability statistics are just one tool in the teachers' toolkit.

Transcribing the writing samples of a whole class would be an arduous task for the classroom teacher. However, if students type their story directly into a computer then only the task of calculating the readability statistics remains. Students could be taught how to do this using *Microsoft Word* inbuilt readability statistics. However, an easier option is to copy and paste written work into an online readability statistics generator, such as *Edit Central*. There is potential for a web developer to create a young writers' user-friendly site by incorporating graphics and levels of challenge. The first level would limit the number of readability statistics generated. Higher levels would provide more extensive, complex readability statistics.

The readability statistics can be used in other ways. For example, teachers could challenge students to increase the number of words they write by keeping a record of word count over a term. Helping students add multi-syllabic words to their vocabulary and their writing could increase the average number of syllables count. Students would need to be trained to note small decimal increases as significant, remembering *The Cat in the Hat* had 0.97 average syllables per word and the complex medical text, where multisyllabic words composed over a quarter of the text, had an average syllable count of 1.87.

In essence, students need help with the complex task of generating clearly written text. Teachers need ways of helping children write better. Used and interpreted appropriately, readability statistics provide objective feedback to students about their written text. Under the guidance of a skilled teacher, students can learn simple ways of improving their writing using readability statistics. The human brain has a superb parsing facility that enables the recognition of sentences, whether full stops are present, or not. Web-based readability formulas are highly sensitive to sentence length and placement of full stops. A

Volume 11, Number 4: November 2010
ISSN: 1535-0975

computer mechanically computes feedback based only on the data present in the document. This incongruous feedback reinforces the importance of placing full stops appropriately and frees both the teacher and the student to work on deeper aspects of creating readable written text.

Volume 11, Number 4: November 2010
ISSN: 1535-0975

Appendix A: Online Readability Calculators

Edit Central
http://www.editcentral.com/gwt1/EditCentral.html
Copy your text and paste into a box. Displays sixteen readability statistics.

Joe's Web Tools
http://www.joeswebtools.com/text/readability-tests/
Copy your text and paste into a box. Displays six readability statistics.

Tyler, S. K. Writing Sample Analyzer (1996-2010)
http://bluecentauri.com/tools/writer/sample.php
Copy your text, up to 5,000 words, and paste into a box. Displays 3 text features and 3 readability statistics.

Child, D. Added Bytes (2003-2009)
http://www.addedbytes.com/code/readability-score/
Copy your text and paste into a box. Displays seven readability statistics.

Readability Analyzer
http://labs.translated.net/text-readability/
Copy your text and paste into a box. This site grades text as easy, average or hard. Provides a list of potentially hard terms.

Simpson, D. The Readability Test Tool (2009-2010)
http://www.read-able.com/
Generates readability feedback about web pages. Copy and paste a URL into the box. Teachers could find this useful in working out what web pages to allocate as reading for their class. Displays six readability statistics and six text property statistics. An explanation of what the results mean follows the statistics.

Juicy Studio (2000-2010)
http://juicystudio.com/services/readability.php
Explains readability statistics. Test the readability of a website by pasting URL into a box. Displays a mixture of readability statistics and text properties. Calculates number of words with 1, 2, 3, or 4 syllables.

Downloadable Readbility Calculator
Source Forge. Net (2007)
http://flesh.sourceforge.net/
This is an open source Java application, available for download, to your computer. Calculates Flesch Reading Ease Score and Flesch-Kincaid Grade Level after pasting text into a box.

Programming Information

Volume 11, Number 4: November 2010
ISSN: 1535-0975

For information on how to code a microcomputer to generate nine sets of readability statistics, see Schuyler, M. R. (1982). A readability formula for use on microcomputers. *Journal of Reading, 25(6),* 560-591. As part of the code, the Dale and Chall frequent words are listed.

References

Baker, S., Gersten, R., & Graham, S. (2003). Teaching expressive writing to students with learning disabilities: Research-based applications and examples. *Journal of Learning Difficulties, 36*(2), 109-123.

Bruning, R. H., Schraw, G. J., Norby, M. M., & Ronning, R. R. (2004). *Cognitive psychology and instruction* (4th ed.). Upper Saddle River, N. J. : Pearson.

Crossley, S. A., Greenfield, J., & McNamara, D. S. (2008). Assessing text readability using cognitively based indices. *TESOL Quarterly, 42*(3), 475-493.

Dale, E., & Chall, J. S. (1948). A formula for predicting readability. *Educational Research Bulletin, 27*(1), 11-28.

Dalecki, L., Lasorsa, D. L., & Lewis, S. C. (2009). The news readability problem. *Journalism Practice, 3*(1), 1 - 12.

Daneman, M., & Stainton, M. (1993). The generation effect in reading and proofreading. *Reading and Writing, 5*(3), 297-313.

Dr Seuss (1957). *The cat in the hat*: Harper Collins.

DuBay, W. H. (2004). *The principles of readability*. Retrieved from http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/1b/bf/46.pdf

Edit Central (2009). *Edit central*. Retrieved from http://www.editcentral.com/gwt1/EditCentral.html

Farr, J. N., Jenkins, J. J., & Paterson, D. G. (1951). Simplification of Flesch reading ease formula. *Journal of Applied Psychology, 35*(5), 333-337.

Flesch, R. F. (1950). Measuring the level of abstraction. *Journal of Applied Psychology, 34*(6), 384-390.

Flesch, R. F. (1951). *How to test readability*. New York: Harper & Brothers.

Fry, E. B. (1968). A readability formula that saves time. *Journal of Reading, 11*(7), 513-578.

Fry, E. B. (1989). Reading formulas: Maligned but valid. *Journal of Reading, 32*(4), 292-297.

Fry, E. B. (2002). Readability versus leveling. *Reading Teacher, 56*(3), 286-291.

Graham, S., & Harris, K. R. (1993). Self-regulated strategy development: Helping students with learning problems develop as writers. *Elementary School Journal, 94*(2), 169-181.

Graham, S., & Harris, K. R. (2000). The role of self-regulation and transcription skills in writing and writing development. *Educational Psychologist, 35*(1), 3-12.

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*(3), 445-476.

Graves, D. H. (1983). *Writing: Teachers and children at work*. Exeter, New Hampshire: Heinemann Educational Books.

Gunning, R. (1968). *The technique of clear writing* (Rev. ed.). New York: McGraw-Hill.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Hillocks, G. (1986). *Research on written composition: New directions for teaching*. New York, N.Y.: National Conference on Research in English.

Klare, G. R. (1984). Readability. In P. D. Pearson (Ed.), *Handbook of Reading Research* (pp. 681-744). New York: Longman.

Lively, B. A., & Pressey, S. L. (1923). A method for measuring the 'vocabulary burden' of textbooks. *Educational Administration and Supervision, 9*, 389-398.

Locke, E. A., & Latham, G. P. (1984). *Goal setting: A motivational technique that works*. Hillsdale, NJ: Prentice Hall.

MacArthur, C. A., Graham, S., & Fitzgerald, J. (Eds.). (2006). *Handbook of writing research*. New York: The Guilford Press.

McLaughlin, G. H. (1969). SMOG grading - a new readability formula. *Journal of Reading, 12*(8), 639-646.

Murray, D. (1984). *Write to learn*. New York: Holt, Rinehart, and Winston.

O'Neill, J. (2000). SMART goals, SMART schools. *Educational Leadership, 57*(5), 46-50.

Pilotti, M., & Chodorow, M. (2009). Error detection/correction in collaborative writing. *Reading and Writing: An Interdisciplinary Journal, 22*(3), 245-260.

Powers, R. D., Sumner, W. A., & Kearl, B. E. (1958). A recalculation of four adult readability formulas. *Journal of Educational Psychology, 49*(2), 99-105.

Queensland Health (2003-2010). *10,000 steps.* Retrieved from http://www.10000steps.org.au/

Rogers, L. A., & Graham, S. (2008). A meta-analysis of single subject design writing intervention research. *Journal of Educational Psychology, 100*(4), 879-906.

Sawyer, R. J., Graham, S., & Harris, K. R. (1992). Direct teaching, strategy instruction, and strategy instruction with explicit self-regulation: Effects on the composition skills and self-efficacy of students with learning disabilities. *Journal of Educational Psychology, 84*(3), 340-352.

Schwartz, H. J. (1980). Teaching stylistic simplicity with a computerized readability formula. *Paper presented at the International Meeting of the American Business Communication Association* (Washington, DC, December 1980). Retrieved from

http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/35/f0/33.pdf

Spache, G. (1953). A new readability formula for primary-grade reading materials. *Elementary School Journal, 53*(7), 410-413.

Sticht, T. G. (1975). *Reading for working: A functional literacy anthology*. Retrieved from http://www.nald.ca/fulltext/sticht/rfw/rfw.pdf

Thorndike, E. L. (1916). An improved scale for measuring ability in reading. *Teachers College Record, 17*, 40-67.

Tracy, B., Reid, R., & Graham, S. (2009). Teaching young students strategies for planning and drafting stories: The impact of self-regulated strategy development. *Journal of Educational Research, 102*(5), 323-332.

Troia, G. A., & Graham, S. (2002). The effectiveness of a highly explicit, teacher-directed strategy instruction routine. *Journal of Learning Disabilities, 35*(4), 290.

Verma, A. K. (2010). New drugs, old drugs. *Medical Journal of Australia, 192*(7), 407-412. Retrieved from http://www.mja.com.au/public/issues/192_07_050410/ver10915_fm.html

Vogel, M., & Washburne, C. (1928). An objective method of determining grade placement of children's reading material. *Elementary School Journal, 28*(5), 373-381.

Waddington, N. J. (2000). *Waddington diagnostic reading and spelling tests 1 & 2*. Strathalbyn, SA: Waddington Educational Resources. Available from http://www.wadd.com.au/