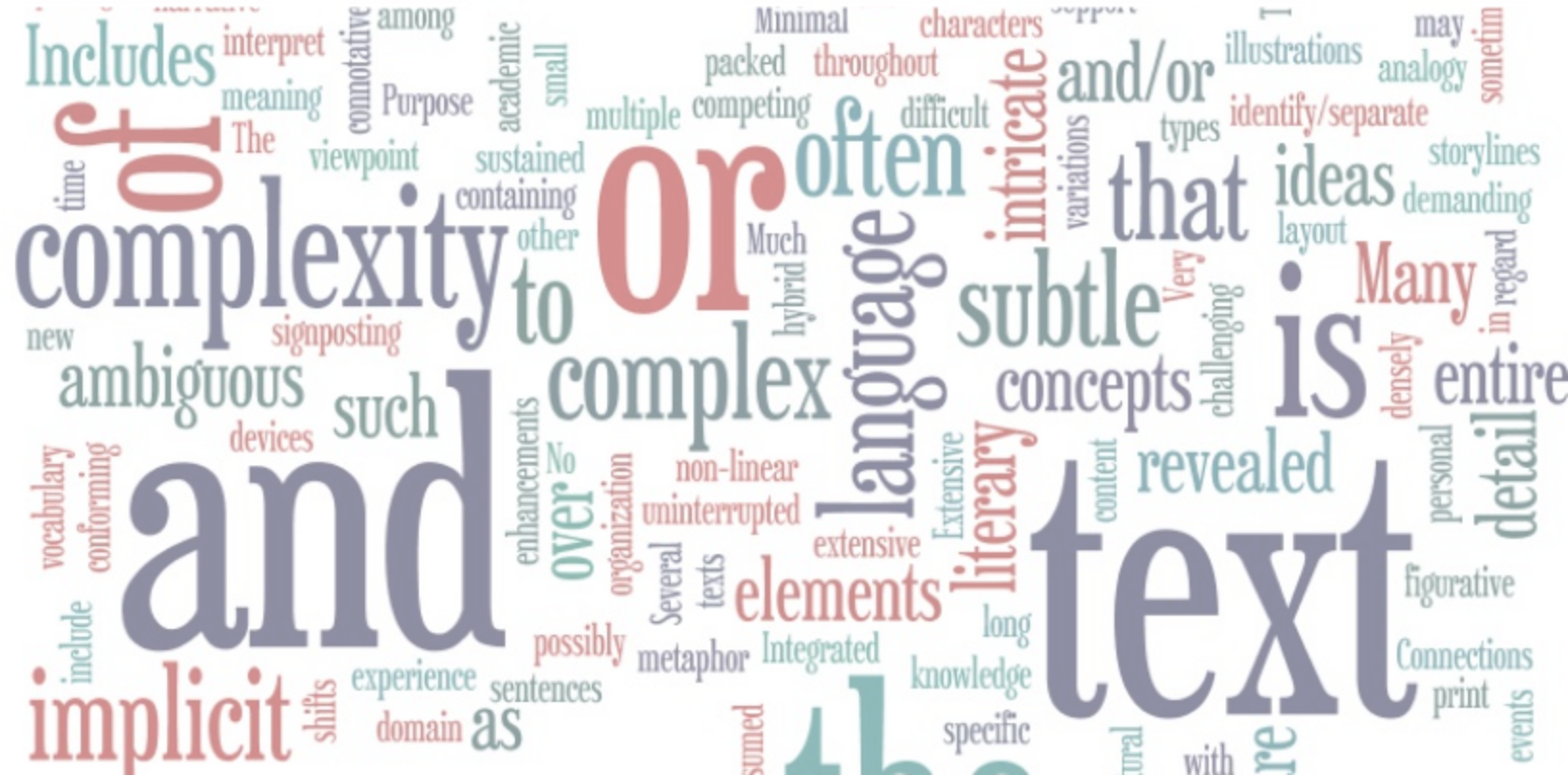
A Beginner's Guide to Text Complexity

AUSSIE, NYCDOE Secondary Literacy Pilot



This document addresses the issues raised around text complexity in the…

Common Core State Standards for English Language Arts & Literacy in History/Social Studies, Science, and Technical Subjects

And in particular…

Standard 10

Range of Reading and Level of Text Complexity

10. Read and comprehend complex literary and informational texts independently and proficiently.

Introduction

Students… "must also develop special skills and strategies for reading text in each of the differing content areas (such as English, science, mathematics and history)— meaning that a student who "naturally" does well in one area may struggle in another."

ACT, Inc. (2006) Reading Between the Lines Iowa City, IA: Author

The Common Core State Standards place an increasing emphasis on getting students to independently read the range and complexity of texts required to be college and career ready. The importance of both increasing the complexity of texts students read and the need for teachers to understand more about what makes their texts challenging arose out of research that showed nearly half of the students who graduate high school need some kind of remediation to cope with the reading required in college and during their careers. The research also showed that the clearest differentiator in reading between students who are college ready and students who are not is the ability to comprehend complex texts.(ACT 2006)

As students move through the grades they are faced with texts that are increasingly longer and more complex in terms of the vocabulary used, sentence structure and text organization. In middle and high schools the texts will present greater conceptual challenges, may include more detailed graphic representation, while at the same time demanding a much greater ability on the part of the reader to synthesize information.

Central to the Standards is the notion that the teacher is able to match students, texts and tasks to promote student learning. Teachers need to know whether students can independently read the range and complexity of grade level discipline-specific materials and if not, what supports and strategies they need. To do this, teachers need to have information on:

- their students as readers;
- the complexity of the texts they are using with the students, i.e. supports and challenges;
- the nature of the tasks they set (how students are going to interact with the text) and the level of support they will provide.

This guide is designed to help teachers to determine the complexity of the texts they use.

What do we mean by text complexity?

Text Complexity is

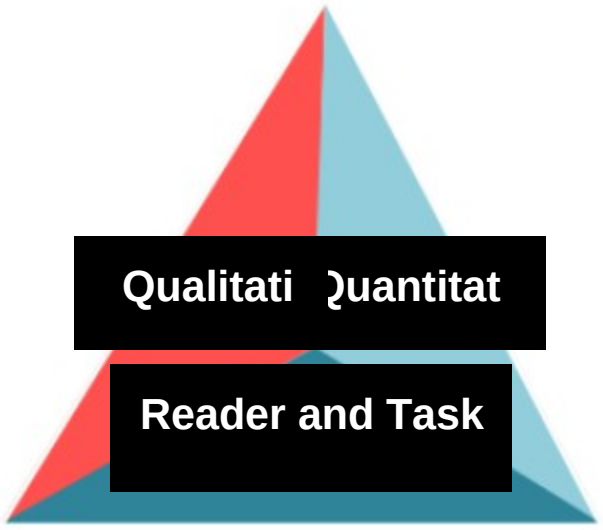
"The inherent difficulty of reading and comprehending a text combined with consideration of reader and task variables; in the Standards, a three-part assessment of text difficulty that pairs qualitative and quantitative measures with reader-task considerations."

CCSS Appendix A

There is no exact science for determining the complexity of a text. Nor is there a single source of information that can accurately summarize the complexity of a text. Teachers need to use their professional judgment as they take into consideration a range of factors.

Three Part Model

The Common Core Standards introduce a three-part model for measuring text complexity. Teachers need to use their professional judgment as they draw on information from all three sources when determining the complexity of a text.



The Standards' Model for Text Complexity

1 Qualitative Measures

The qualitative measures of text complexity require an informed judgment on the difficulty by considering a range of factors. The Standards use purpose or levels of meaning, structure, language conventionality and clarity and the knowledge demands as measures of text difficulty. (pg 6, CCSS Appendix A)

2 Quantitative Measures

Quantitative measures of text complexity use factors such as sentence and word length and frequency of unfamiliar words to calculate the difficulty of the text and assign a single measure (grade level equivalent, number, Lexile etc). There are many formulas for calculating text difficulty and, while they provide a guide, the readability or difficulty level of a text can vary depending on which formulas or measures are used. (pg 8, CCSS Appendix A)

3 Reader and Task

The third measure looks at what the student brings to the text and the tasks assigned. Teachers need to use their knowledge of their students and the texts to match texts to particular students and tasks. (pg 9, CCSS Appendix A)

How do we determine the complexity of texts?

The Common Core Standards raise the expectation for students in terms of the complexity of the texts they read. This means teachers need to be familiar with the level of complexity expected at the grade levels they teach and how these compare to the complexity of the texts they use in their classes. Several considerations should guide teachers in selecting their texts and designing text-based instruction.

1. Quantitative Measures

The quantitative measures provide a very useful guide in determining the complexity of texts. They are, however, not sufficient when used in isolation. Most publishers give grade band equivalents, or Lexile levels, for their texts. A book with a Lexile of 1200 will be considerably more complex than one with a 770 Lexile. The quantitative measure indicates how complex a text is, but does not explain the nature of the complexity. Quantitative measures are determined using readability formulas.

Readability Formulas

There are five readability formulas that are commonly used to measure the complexity of texts. While all can be calculated manually, there are computer programs that calculate readability when you paste in a section of 100-200 words. For instance, lexile.com, offers a free readability analysis using the Lexile framework, and provides results that are aligned to the Common Core State Standards. Other commonly used readability formulas include:

The Flesch Reading Ease Readability Formula and the The Flesch-Kincaid Grade Level Readability Formula calculate difficulty using sentence length and number of syllables per word.

Gunning's Fog Index (or FOG) Readability Formula uses sentence length and percentage of Foggy words (words with three or more syllables).

The Dale-Chall Readability Formula uses sentence length and percentage of difficult words (words that do not appear on the familiar word list).

The quantitative measures are not sufficient in themselves. There is no question that sentence length, number of syllables in words, and word frequency impact difficulty. There is a wide range of factors that influence the complexity of a text. To be able to match students and texts, teachers need to know where the complexity lies within a text to ensure they provide students with the supports and strategies needed to successfully read the texts.

See appendix A of this guide for more detail on readability formulas.

2. Qualitative factors for describing complexity

Qualitative measures of text complexity provide valuable information when making decisions about the complexity of the text and how it could best be used with students. The Common Core State Standards identify a range of qualitative factors that interact to contribute to the overall complexity. Rubrics have been developed for both literary and informational texts that include descriptors for:

- layout;
- purpose and meaning;
- text structure;
- language features;
- knowledge demands.

Not all descriptors described in each of the categories will necessarily occur together at each level of complexity. A text may have very simple vocabulary and short, simple sentences yet still be complex because the ideas expressed are subtle.

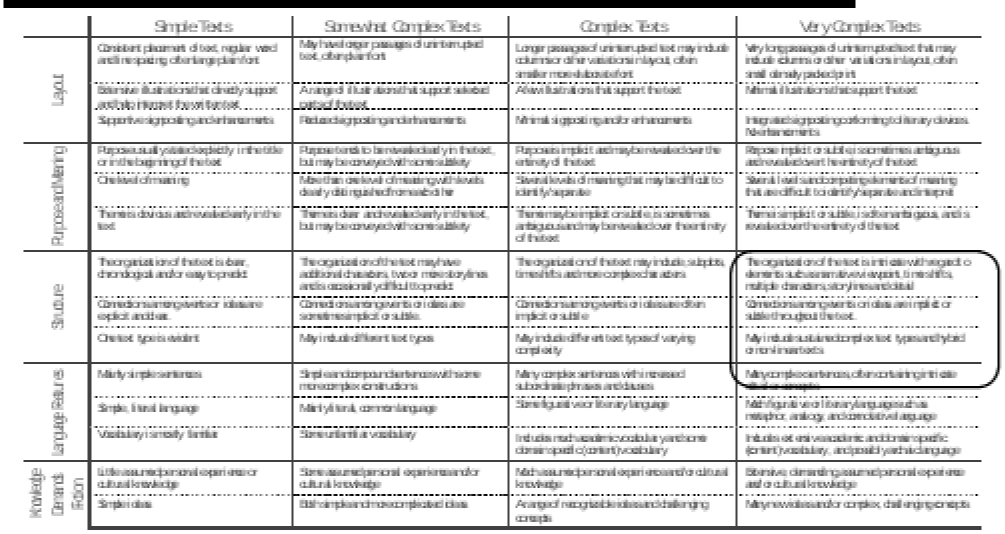
The organization of the text is intricate with regard to elements such as narrative viewpoint, time shifts, multiple characters, storylines and detail

Connections among events or ideas are implicit or subtle throughout the text

Includes sustained complex text types and hybrid or non-linear texts

Many complex sentences, often containing intricate detail or concepts

[redacted]



By using a rubric it is easy to see where the complexity of the text lies. If teachers know what aspects of the text are likely to be challenging for students, they can make decisions about the suitability of a text and what strategies or supports students may need to read it successfully.

(Please note that this rubric is available on the Common Core Library.)

3. What about the reader and the task?

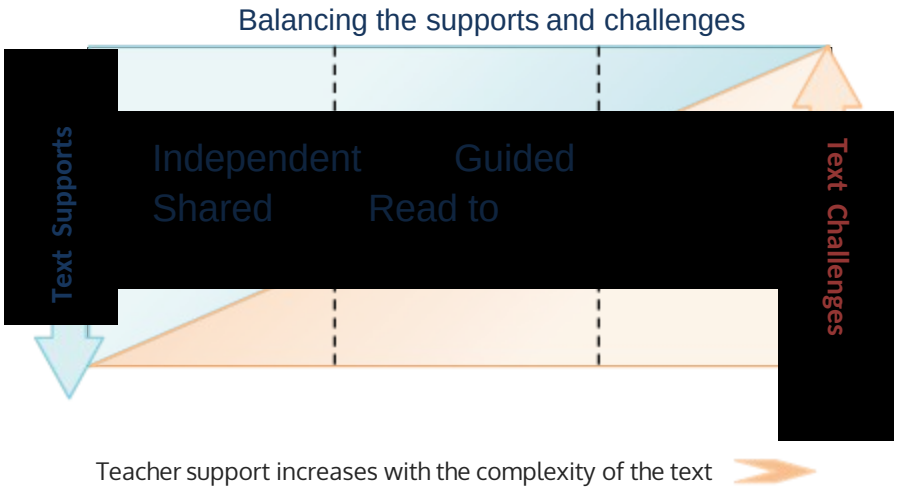
Qualitative and quantitative measures describe the complexity within the text. The Common Core's definition of text complexity also encourages educators to consider our students and how we are expecting them to interact with the text. In any class there will be a range in the students' ability to read complex texts. Teachers will need to use their professional judgment when making decisions about what texts to use and how they should be used. This professional judgment is dependent on the teachers':

- knowledge of their students as readers;
- understanding of the complexity of the texts;
- ability to use a range of instructional approaches flexibly.

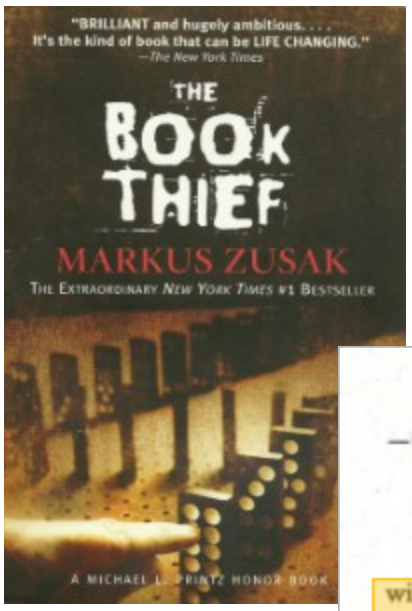
Tasks, like texts, become more complex as students consider ideas and information in different ways. There is a "gear shift" from locating and evaluating items of information on a topic through to locating, evaluating, and synthesizing information from several different sources.

Balancing the supports and challenges

The more complex the text, the more support students will need. Students will be introduced to increasingly complex texts throughout middle and high school. This is done through a gradual release of responsibility where complex texts are introduced in a supportive context that facilitates higher levels of independence.



Reading complex texts requires students to actively engage with texts as they make meaning. This requires commitment and risk taking on the part of the reader. Students will not put this amount of effort into texts that are irrelevant and uninteresting or where they see no value in the tasks assigned.



Example of Text Complexity, Literary:

The Book Thief by Markus Zusak

Exemplar Text for Grades 9-10 Text Complexity Band (Appendix B CCSS)
Lexile: 730L

—Of course, an introduction.

A beginning.

Where are my manners?

I could introduce myself properly, but it's not really necessary. You will know me well enough and soon enough, depending on a diverse range of variables. It suffices to say that at some point in time, I will be standing over you, as genially as possible. Your soul will be in my arms. A color will be perched on my shoulder. I will carry you gently away.

At that moment, you will be lying there (I rarely find people standing up). You will be caked in your own body. There might be a discovery; a scream will dribble down the air. The only sound I'll hear after that will be my own breathing, and the sound of the smell, of my footsteps.

The question is, what color will everything be at that moment when I come for you? What will the sky be saying?

Personally, I like a chocolate-colored sky. Dark, dark chocolate. People say it suits me. I do, however, try to enjoy every color I see—the whole spectrum. A billion or so flavors, none of them quite the same, and a sky to slowly suck on. It takes the edge off the stress. It helps me relax.

★ ★ ★ A SMALL THEORY ★ ★ ★

People observe the colors of a day only at its beginnings and ends, but to me it's quite clear that a day merges through a multitude of shades and intonations, with each passing moment. A single *hour* can consist of thousands of different colors. Waxy yellows, cloud-spat blues. Murky darknesses. In my line of work, I make it a point to notice them.

As I've been alluding to, my one saving grace is distraction. It keeps me sane. It helps me cope, considering the length of time I've been

The narrator is not identified until part way through the text

The personification of death throughout the text

Much of the text is figurative with extensive use of metaphor

Figurative language

Much of the text is figurative with extensive use of metaphor

Innovative stylistic techniques are used. The most obvious is the narrator's (Death) use of boldface text to relay certain information.

The Book Thief, continued

This text highlights the importance of not relying solely on quantitative measures. The Lexile level is 730 L, which would suggest the book is suitable for grades 3 and 4 students (see chart on pg 14 of this guide). However, it is an exemplar text for grades 9-10 in the Standards. The complexity becomes evident when the qualitative measures are used.

Students are likely to find the following characteristics challenging:

- the historical setting;
- much of the text is figurative with extensive use of metaphor, including the personification of death itself;
- the text is long – 552 pages;
- the innovative stylistic techniques used. The most obvious is the narrator Death's use of boldface text to relay certain information;
- the intertwining, multiple themes.

	Simple Texts	Somewhat Complex Texts	Complex Texts	Very Complex Texts
Layout	Consistent placement of text, regular word and line spacing, often large plain font	May have longer passages of uninterrupted text, often plain font	Longer passages of uninterrupted text may include columns or other variations in layout, often smaller more elaborate font	Very long passages of uninterrupted text that may include columns or other variations in layout, often small densely packed print
	Extensive illustrations that directly support and help interpret the written text	A range of illustrations that support selected parts of the text	A few illustrations that support the text	Minimal illustrations that support the text
	Supportive signposting and enhancements	Reduced signposting and enhancements	Minimal signposting and/or enhancements	Integrated signposting conforming to literary devices. No enhancements
Purpose and Meaning	Purpose usually stated explicitly in the title or in the beginning of the text	Purpose tends to be revealed early in the text, but may be conveyed with some subtlety	Purpose is implicit and may be revealed over the entirety of the text	Purpose implicit or subtle, is sometimes ambiguous and revealed over the entirety of the text
	One level of meaning	More than one level of meaning, with levels clearly distinguished from each other	Several levels of meaning that may be difficult to identify/separate	Several levels and competing elements of meaning that are difficult to identify/separate and interpret
	Theme is obvious and revealed early in the text	Theme is clear and revealed early in the text, but may be conveyed with some subtlety	Theme may be implicit or subtle, is sometimes ambiguous and may be revealed over the entirety of the text	Theme is implicit or subtle, is often ambiguous, and is revealed over the entirety of the text
Structure	The organization of the text is clear, chronological and/or easy to predict	The organization of the text may have additional characters, two or more storylines and is occasionally difficult to predict	The organization of the text may include, subplots, time shifts and more complex characters	The organization of the text is intricate with regard to elements such as narrative viewpoint, time shifts, multiple characters, storylines and detail
	Connections between events or ideas are explicit and clear.	Connections among events or ideas are sometimes implicit or subtle.	Connections among events or ideas are often implicit or subtle	Connections among events or ideas are implicit or subtle throughout the text.
	One text type is evident	Includes different text types	Includes different text types of varying complexity	Includes sustained complex text types and hybrid or non-linear texts
Language	Mainly simple sentences	Simple and compound sentences with some more complex constructions	Many complex sentences with increased subordinate phrases and clauses	Many complex sentences, often containing intricate detail or concepts

	Simple, literal language	Mainly literal, common language	Some figurative or literary language	Much figurative or literary language such as metaphor, analogy, and connotative language
	Vocabulary is mostly familiar	Some unfamiliar vocabulary	Includes much new vocabulary and some domain specific (content) vocabulary	Includes extensive unfamiliar vocabulary, and possibly archaic language
Knowledge Demands Fiction	Little assumed personal experience or cultural knowledge	Some assumed personal experience and/or cultural knowledge	Much assumed personal experience and/or cultural knowledge	Extensive, demanding, assumed personal experience and/or cultural knowledge
	Simple ideas	Both simple and more complicated ideas	A range of recognizable ideas and challenging concepts	Many new ideas and/or complex, challenging concepts

Text Complexity Rubric for The Book Thief

Example of Text Complexity, Informational:

"Thinking about physics while scared to death (on a falling roller coaster)" by Jearl Walker

Roundabout: Readings from the Amateur Scientist in Scientific American. New York: Scientific American, 1985.

Challenging abstract concepts

domain-specific vocabulary

nominalization

Many complex sentences with increased subordinate phrases and clauses or transition words

domain-specific vocabulary

Many complex sentences with increased subordinate phrases and clauses or transition words

THE AMATEUR SCIENTIST

Thinking about physics while scared to death (on a falling roller coaster)

by Jearl Walker

The rides in an amusement park not only are fun but also demonstrate principles of physics. Among them are rotational dynamics and energy conversion. I have been exploring the rides at Geauga Lake Amusement Park near Cleveland and have found that nearly every ride offers a memorable lesson.

To me the scariest rides at the park are the roller coasters. The Big Dipper is similar to many of the roller coasters that have thrilled passengers for most of this century. The cars are pulled by chain to the top of the highest hill along the track. Released from the chain as the front car begins its descent, the unpowered cars have almost no speed and only a small acceleration. As more cars get onto the downward slope the acceleration increases. It peaks when all the cars are headed downward. The peak value is the product of the acceleration generated by gravity and the sine of the slope of the track. A steeper descent generates a greater acceleration, but packing the coaster with heavier passengers does not.

When the coaster reaches the bottom of the valley and starts up the next hill, there is an instant when the cars are symmetrically distributed in the valley. The acceleration is zero. As more cars ascend, the coaster begins to slow, reaching its lowest speed just as it is symmetrically positioned at the top of the hill.

A roller coaster functions by means of transfers of energy. When the chain hauls the cars to the top of the first hill, it does work on the cars, endowing them with gravitational potential energy, the energy of a body in a gravitational field with respect to the distance of the body from some reference level such as the ground. As the cars descend into the first valley much of the stored energy is transferred into kinetic energy, the energy of motion.

If the loss of energy to friction and air drag is small, the total of the potential and kinetic energies must remain constant throughout the descent and even throughout the rest of the ride. The coaster gains kinetic energy and speed at the expense of potential energy. If the first valley is at ground level, the transfer is complete, and for a moment the coaster has all its energy in the form of kinetic energy.

Without energy losses the coaster could climb any number of hills as high as the one from which it is released (but no higher). To be sure, friction and air drag do remove energy from the coaster, and its total energy content dwindles. It can no longer climb high hills, which is why the last stages of the track consist only of low hills.

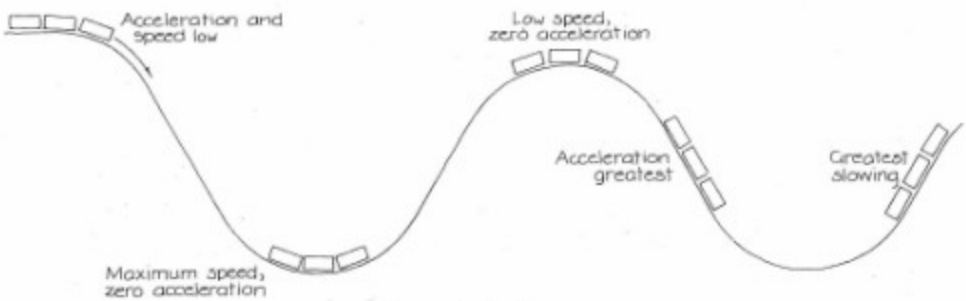
The length of a ride on a roller coaster depends on the speed. If the ride is to be fast, the launching hill should be high so that the total energy is large. The rest of the track should be low so that most of the energy remains kinetic.

The choice of a seat on a roller coaster makes a difference in the ride. Some people prefer the front seat because the descent from the launching site presents the pleasingly frightening illusion of falling over the edge of a cliff. Other people prefer the psychological security of the rear seat.

The choice of a seat also determines the forces felt by the passenger. Consider the first descent. The front car starts down slowly because little of the coaster's energy is then kinetic. The speed of the cars increases as an exponential function of time, so that the rear car starts down at a much higher speed than the front car did. Although the passengers in the front car get an unobstructed view of the descent, the passengers in the rear car have a stronger sense of being hurled over the edge.

At the edge one force on the passenger is from the change in the direction of his momentum vector. Initially the vector is horizontal, but soon it points toward the valley. The force necessary to effect this change in direction is delivered by the safety bar or seat belt that keeps the passenger in the car. That force, which points downward and back toward the hill, is part of the thrill of the ride. A passenger in the rear feels the force more than a passenger in the front because the size of the force is proportional to the momentum, which is greater for the passenger in the rear.

The story is different in the valley. Again a force from the coaster is necessary to redirect the passenger's momentum. This time the momentum is initially downward toward the bottom of the valley and then is redirected toward the top of the next hill. The front passenger has a large momentum and is subjected



The energetics of a roller coaster

"Thinking about physics while scared to death (on a falling roller coaster)" by Jearl Walker, continued

Exemplar Text for Grades 9-10 Text Complexity Band (Appendix B)
Flesch-Kincaid Grade Level puts the readability at the 8th grade level.

The language structure of this text is relatively straightforward; however, the complexity lies in the domain-specific vocabulary, complex embedded sentences and the difficulty of the ideas being explained.

Students are likely to find the following characteristics challenging:

- the domain-specific vocabulary;
- the knowledge demands of the physics concepts of motion and force;
- small, densely packed print;
- nominalization;
- minimal use of diagrams, e.g. to show directions of forces.

	Simple Texts	Somewhat Complex Texts	Complex Texts	Very Complex Texts
Layout	Consistent placement of text, regular word and line spacing, often large plain font	May have longer passages of uninterrupted text, often plain font	Longer passages of uninterrupted text may include columns or other variations in layout, often smaller more elaborate font	Very long passages of uninterrupted text that may include columns or other variations in layout, often small densely packed print
	Graphics and pictures that directly support and help interpret the written text	Graphs, pictures, tables, charts that directly support the text	Essential integrated graphics, tables, charts, formula (necessary to make meaning of text)	Extensive, intricate, essential integrated tables, charts, formulas necessary to make meaning of text
	Simple indexes, glossaries	Indexes, glossaries, occasional quotes, references	Quotes, concluding appendices, indexes, glossaries, bibliography	Abstracts, footnotes, citations and detailed indexes, appendices, bibliography
	Supportive signposting and enhancements	Reduced signposting and enhancements	Minimal signposting and/or enhancements	Integrated signposting conforming to disciplinary formats. No enhancements
Purpose and Meaning	A single or simple purpose conveying clear or factual information	Purpose involves conveying a range of more detailed information	Purpose includes explaining or interpreting information	Purpose may include examining/evaluating complex, sometimes theoretical and contested information
	Meaning is clear, concrete with a narrow focus	Meaning is more involved with a broader focus	Meaning includes more complex concepts and a higher level of detail	Meaning is intricate, with abstract theoretical elements
Structure	The organization of the text is clear or chronological and/or easy to predict	The organization of the text may include a thesis or reasoned explanation in addition to facts	The organization of the text may contain multiple pathways, more than one thesis and/or several genres	The organization of the text is intricate or specialized for a particular discipline
	Connections between ideas, processes or events are explicit and clear.	Connections between some ideas, processes or events are implicit or subtle	Connections between an expanded range ideas, processes or events are deeper and often implicit or subtle.	Connections between an extensive range ideas, processes or events are deep, intricate and often implicit or subtle.
	One text type is evident	Includes different text types	Includes different text types of varying complexity	Includes sustained complex text types and/or specialized, hybrid text types
Language	Mainly simple sentences	Simple and compound sentences with some more complex constructions	Many complex sentences with increased subordinate phrases and clauses or transition words	Mainly complex sentences, often containing multiple concepts

	Simple language style, sometimes with narrative elements	Increased objective style and passive constructions with higher factual content	Objective/passive style with higher conceptual content and increasing nominalization	Specialized disciplinary style with dense conceptual content and high nominalization
	Vocabulary is mostly familiar	Vocabulary includes some unfamiliar, context-dependent words	Includes much academic vocabulary and some domain specific (content) vocabulary	Includes extensive academic and domain specific (content) vocabulary
Knowledge Demands Informational	General topic is familiar, with details known by reader	General topic is familiar, with some details new to reader	General topic is somewhat familiar but with many details unknown to reader	General topic is mostly unfamiliar with most details unknown to reader
	Simple, concrete ideas	Both simple and more complicated, abstract ideas	A range of recognizable ideas and challenging abstract concepts	Many new ideas and/or complex, challenging, abstract and theoretical concepts

Text Complexity Rubric for “Thinking about physics while scared to death”

Professional Practice

The way in which ideas and information are presented to students and the opportunities and scaffolds provided for them to engage with texts are critical elements of teacher practice that can shape students' success in navigating complex texts.

Teachers need to provide multiple,

Planning for Support

Tasks, like texts, become more complex as students think about ideas and information in different ways. When considering the complexity of the text teachers need to take into account the tasks they set, as well as their knowledge of their students as readers.

When introducing texts teachers need to consider the challenges in the text and the strategies students need.

Group	10th Grade ELA	10th Grade Physics
Text Title	**The Book Thief by Markus Zusak**	**Flying Circus of Physics by Jearl Walker**
Complexity Level	Exemplar Text for Grades 9-10 Text Complexity Band (Appendix B CCSS) The Lexile level is 730 L	Exemplar Text for Grades 9-10 Text Complexity Band (Appendix B CCSS) Flesch-Kincaid Grade Level puts the readability at the 8th grade level
Texts Supports	Familiar vocabulary.	Links to well known phenomenon
Text Structure and Concepts Challenges	1. figurative language with extensive use of metaphor and analogy 2. the text is a very long 552 pages 3. the use of innovative stylistic techniques. The most obvious is narrator Death's use of boldface text to relay certain information	1. the domain specific vocabulary- 2. the knowledge demands around motion and force 3. difficulty of the concepts being explained 4. lack of illustrations or diagrams to help visualize the explanations.
Planned Teacher Supports	Activate connections to 8th grade work on Holocaust Read first 3 chapters aloud to familiarize students with language Guide thinking around narrator Provide activities that locate figurative language Character webs to track changes in characters	Make links to previous learning Use anticipation guides to predict. Sketch to stretch to help visualize concepts being introduced.
:		

Appendix A (of "Beginner's Guide to Text Complexity")

Readability Formulas

There are five commonly used readability formulas. While all can be calculated manually, there are computer prgrams that do this for you by pasting in a section of 100-200 words. ©ReadabilityFormulas.com , for instance, offers a free readability analysis and gives results using the four most common used readability scores.

1. The Flesch Reading Ease Readability Formula
The specific mathematical formula is:

Readability Ease = 206.835 – (1.015 x average sentence length) – (84.6 x average syllable per word).

The higher the number, the easier the passage: 90-100, very easy and 0-29, very confusing.

2. The Flesch-Kincaid Grade Level Readability Formula
The specific mathematical formula is:

Flesch-Kincaid Reading Age = (0.39 x average sentence length) + (11.8 x average syllable per word) - 15.59

The score is given as a grade level equivalent – a score of 9.4 would be 9th grade. Flesch-Kincaid Grade Level Readability Formula is built into the MS-Word application.

3. Gunning's Fog Index (or FOG) Readability Formula
The specific mathematical formula is:

The FOG formula takes into consideration (1) the total number of words, (2) the number of words of three or more syllables, and (3) the total number of sentences.

Grade Level = 0.4 (Average Sentence Length + Percentage of Hard Words).

Hard words are defined as words with three or more syllables.

4. The Dale-Chall Readability Formula
The specific mathematical formula is

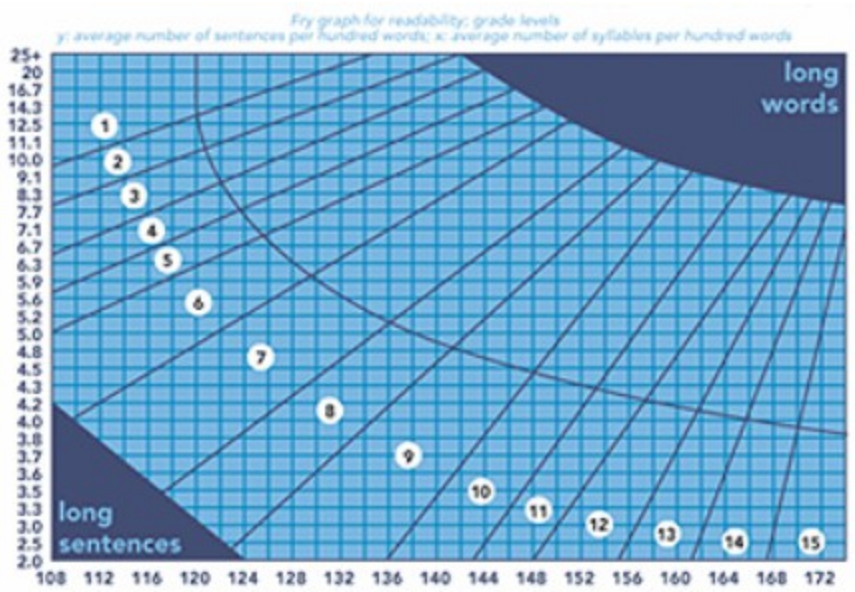
Raw Score = 0.1579 Percentage of Difficult Words + 0.0496 Average Sentence Length in words + 3.6365

Raw scores convert to grade levels with scores 5.0 to 5.9 = Grades 5 – 6 and 8.0 to 8.9 = Grades 11 - 12

Difficult words are words that do not appear on the Chall word list of over 3000 familiar words.

5. The Fry Graph Readability Formula

The Fry readability formula uses three 100 word passages and calculates the average number of words per sentence and the number of syllables per 100



words.

.

Scores that appear in the dark area (long sentences and long words) are invalid.

Many companies are in the process of adjusting their leveling systems to align to the raised expectations for standards. The following chart shows how MetaMetrics has realigned its Lexile ranges to align to expectations for College and Career Readiness.

Fig 3: Text Complexity Grade Bands and Associated Lexile Ranges (in Lexiles) – from page 8 of the Common Core State Standards, Appendix A

Text Complexity Grade Band in the Standards	Old Lexile Ranges	Lexile Ranges Aligned to CCR
K–1	N/A	N/A
2–3	450–725	450–790
4–5	645–845	770–980
6–8	860–1010	955–1155

Reader and Task

To use the Lexile framework online go to: http://www.lexile.com/

To find the readability formula for a text there are two options.

--Use the "Quick Book Search" (most suitable for published books)

--Use the "Lexile Analyzer" (Most suitable for articles. A free account is required to use this function.)