

Patrick Cadwell

## Readability and Controlled Language:

Does the study of readability have merit in the field of controlled language, and is readability increased by applying controlled-language rules to texts?

A dissertation submitted to Dublin City University in partial fulfilment of the requirements for the degree of MA in Translation Studies.

School of Applied Language and Intercultural Studies

Submission: September 2008

# Declaration

*'I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of MA in Translation Studies, is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.'*

Signed: \_\_\_\_\_ StudentNumber: \_\_\_\_\_  
(Candidate)

Date: \_\_\_\_\_

# Table of Contents

---

Abstract .....	5
Acknowledgments.....	6
1.0 Introduction.....	7
2.0 Literature Review.....	8
2.1 Motivations .....	8
2.1.1 Need for empirical research.....	8
2.1.2 Assumptions held in the literature .....	8
2.1.3 Lack of terminological rigour .....	9
2.1.4 Studies that inspired this experiment .....	9
2.2 Controlled Language (CL).....	10
2.2.1 Definition .....	10
2.2.2 Different types of CL.....	10
2.2.3 Best-known examples of CL.....	11
2.2.4 Advantages and disadvantages .....	11
2.2.5 Summary .....	11
2.3 Readability .....	12
2.3.1 Defining and measuring readability.....	12
2.3.2 Predicting readability: readability formulas.....	13
2.3.3 Producing readability: the CL created for this experiment.....	14
2.3.4 Extra-linguistic variables .....	18
2.3.5 Summary .....	19
2.4 Other metrics.....	19
2.4.1 Comprehensibility.....	19
2.4.2 Translatability .....	19
2.4.3 Usability.....	19
2.4.4 Others.....	20
2.5 Introducing terminological rigour.....	20

<b>3.0 Methodology .....</b>	<b>23</b>
3.1 Outline of the questions to be answered in the experiment .....	23
3.2 Theoretical framework.....	24
3.3 Textual data on which the survey was based.....	25
3.3.1 Selecting appropriate texts to use as NCL versions.....	25
3.3.2 Isolating different levels of readability .....	26
3.3.3 Creating CL versions .....	27
3.4 Survey .....	30
3.4.1 The sample of participants.....	30
3.4.2 The time-lag between Stage 1 and Stage 2 .....	32
3.5 The questionnaire.....	33
3.6 Limitations .....	38
3.7 Methodological points to avoid .....	38
<b>4.0 Data Analysis .....</b>	<b>40</b>
4.1 Would the CL version be preferred by readers?.....	40
4.2 Would the other two pillars of readability be altered in the CL version?.....	41
4.3 Would the formulas' predictions correspond to readers' opinions?.....	44
4.4 Would extra-linguistic variables impact on readability? .....	45
4.5 Would certain parts of the texts prove to be more difficult than others?.....	48
4.6 Surprising extra data .....	49
<b>5.0 Conclusion .....</b>	<b>50</b>
<b>Reference List .....</b>	<b>52</b>

## Appendices

# Abstract

---

**Author:** Patrick Cadwell

**Title: Readability and Controlled Language:** does the study of readability have merit in the field of controlled language, and is readability increased by applying controlled-language rules to texts?

This study conducted a survey of text-user attitudes to discover: whether the concept of readability had merit in the field of controlled language; and whether readability increased by applying controlled-language rules to a sample of technical texts. This experiment was an attempt to provide much-needed empirical data to a neglected area of controlled-language research, and to examine the concept of readability that appears to be misunderstood, undervalued and misused. This experiment was carried out in two stages: in one stage, participants in two groups (one with domain expertise and one control) were asked to read a controlled version of a technical passage and then complete a questionnaire. Then in another stage, the same participants were required to read an uncontrolled version of the same passages and fill out a questionnaire. In this way, it was possible to examine attitude variance depending on whether a controlled or uncontrolled version of the text had been read. This study found that a majority of participants in both groups determined the controlled versions to be more readable. The participant samples in this experiment were too small to be generalised to larger populations. However, working as a pilot study, the trends identified in this paper indicate useful methodological recommendations for future research: in terms of readability formula selection; in terms of readability testing methodology; and in terms of conceptual mapping in the field of controlled language.

**Keywords:** readability; controlled language; reader preference; empirical data.

# Acknowledgments

---

I wish to thank Dr Sharon O'Brien for introducing me to the topic of readability, for helping me to gain access to the many Symantec resources used in this experiment, and for guiding me through the dissertation process. Her assistance has been invaluable.

I also wish to thank Dr Fred Hollowood, Dr Johann Roturier and Nora Aranberri of Symantec for their help and encouragement throughout the experiment. In particular, the access to corporate resources that Dr Hollowood permitted me was much appreciated.

Finally, I thank all the participants, both in the Symantec group and in the Control group, who took time out of their busy schedules to help me with my survey. I am indebted to them all.

# 1.0 Introduction

---

This study examined the concepts of readability and controlled language (CL) by conducting a survey of text-user attitudes. Much work in the literature has experimented on related concepts, such as translatability or comprehensibility. However, few empirical studies exist which explore readability and its relationship to CL. Moreover, though readability is regularly discussed by CL researchers, it appears to be somewhat misunderstood by many. It is hoped that this study will identify the important issues to consider when examining how CL affects readability, and will illustrate the usefulness of readability theory to understanding document production.

The paper begins, in Chapter 2, with a literature review. Here, the motivations behind the work are explained, and the field of CL is introduced. Then, the theory behind the concept of readability, and the complex interplay between it and other forms of CL text analysis are outlined. Chapter 3 presents the methodology behind the survey carried out in this experiment: the questions it hoped to answer; the theoretical framework on which it was based; how and why the data were gathered; and the limitations to these data and the methodology itself. Chapter 4 gives a detailed analysis of the answers to the questions posed at the outset of the experiment. Finally, Chapter 5 moves on to the conclusions that can be drawn from the data.

## 2.0 Literature Review

---

This chapter aims to show that the concept of readability is a useful one in the field of CL. It explains the complex relationship between readability and other forms of text analysis and the influence of this relationship on the present survey.

### *2.1 Motivations*

Three major factors motivated this study:

#### **2.1.1 Need for empirical research**

There appears to be a lack of empirical data in the field of CL: Knops (2000 p134) points out that:

Generally speaking, there is an urgent need for facts and figures obtained in experimental situations and real-life production environments and relating to the effects of particular CL standards, rules and rule sets on readability and translatability.

In recent years, several empirical studies have been published in the field of CL rules: they have focused largely on more machine-oriented topics, such as translatability, comprehensibility, proof-reading effort, etc. More human-oriented analyses – such as readability, usability, etc. - have been neglected. According to Hayes, Maxwell and Schmandt (1996 pp84-85), this may be because readability advantages are harder to quantify than those of translatability, comprehensibility, etc.

#### **2.1.2 Assumptions held in the literature**

Several under-tested and unchallenged assumptions about readability now hold firm in the literature. For example, “reducing the complexity of syntactic structures of a text increases its readability” (Spaggiari, Beaujard and Canneson 2003 p152). Similarly, Reuther (1998 p174) claims that:

It is a well known and indisputable fact within the CL community that the use of a Controlled Language (CL) in technical documentation leads to quality improvement with respect to readability, consistency and translatability.

In the literature, it is difficult enough to find a consistent definition for readability, let alone empirical evidence for such strong statements. This paper intends to show the ultimate value of readability in the field of CL: regardless of how easy a CL makes a document to translate or comprehend, these benefits will be for nothing if the text is written in a way that causes the reader to discard it.

### **2.1.3 Lack of terminological rigour**

The literature contains an abundance of terms used to describe how CL texts can be analysed and evaluated (hereafter referred to as metrics). Metrics in the field of CL include: readability; comprehensibility; translatability; usability; post-editing effort; consistency, legibility; acceptability; accessibility; learnability. Depending on the author, these terms can be treated as sharing many, all or no characteristics. Clearly, there is a need for a more systematic and rigorous definition and treatment of these concepts.

Before introducing the theoretical issues that will be dealt with in this study, it is important to highlight the studies that inspired this work.

### **2.1.4 Studies that inspired this experiment**

The methodologies and references in the following works constituted a useful starting point for the present study: Shubert et al. (1995) provided empirical evidence of readability being improved with the application of Simplified English rules; Hayes, Maxwell and Schmandt (1996) showed that CL versions created by them were preferred by readers; finally, Moller (2003) examined the preferences of real users for texts controlled with Simplified English rules.

Having established that there was some need for this study, let us now look at the theoretical issues involved in it.

## ***2.2 Controlled Language (CL)***

This section defines a CL, outlines the different types of CL that exist, and explains the advantages and disadvantages of applying a CL.

### **2.2.1 Definition**

A CL forms part of the document development chain, usually where large volumes of documentation are created by multiple authors. It is:

...an explicitly defined restriction of a natural language that specifies constraints on lexicon, grammar, and style. The overall aim here is the reduction in ambiguity, redundancy, size and complexity (Huijsen 1998 p2).

CL is often applied to technical texts in the commercial domain because natural language tends to be structurally complex and potentially ambiguous (Mitamura et al. 2003 p87). By controlling language, "...reading texts and performing associated tasks can be more efficient and effective, and computational text processing more reliable (Huijsen 1998 p1)". A CL should not have a major negative impact on author productivity and must still be acceptable to its readers; otherwise, its application will be unsuccessful (Hayes, Maxwell and Schmandt 1996 p86).

### **2.2.2 Different types of CL**

A CL can be categorised by its intended use: if the CL is designed to improve the document for human readers, especially non-native speakers, then it is said to be a human-oriented controlled language (HOCL); however, if it is designed to improve the computational processing of a text, then it is said to be a machine-oriented controlled language (MOCL) (Reuther 2003 p124). In general, then:

A HOCL's objective is to improve readability and comprehensibility whereas the primary objective of a MOCL is to improve translatability (O'Brien 2003 p105).

### **2.2.3 Best-known examples of CL**

Roturier (2006 pp47-48) outlines the development of the most significant CLs over time. Starting with Ogden's Basic English in the 1930s, there has been: Caterpillar Fundamental English dating back to the 1960s; the Plain English Program launched in the 1970s (see Crystal 1997 pp382-383 for an introduction); AECMA Simplified English (hereafter SE) dating back to the 1980s (see Wojcik, Holmback and Hoard 1998); and, most recently, Global English (see Means, Chapman and Liu 2000 p8). SE is perhaps the best known example of a modern CL; it is of particular interest to this study because it is the only major 'HOCL', and because it was originally instigated to improve readability (Wojcik, Holmback and Hoard 1998 p114).

### **2.2.4 Advantages and disadvantages**

The main advantages claimed of CL are that: it makes texts easier to manipulate by humans and computers; it allows tasks involving the texts to be performed better; it makes texts more consistent and easier to maintain (Huijsen 1998 p3).

There are equally many disadvantages to applying CL: its implementation is expensive and complex; it can be difficult to learn and can be met with reluctance by authors who find their creativity stifled; its rules can conflict, be contradictory and open to different interpretations; and it can be difficult to evaluate the effect of individual rules because so many variables are involved in the production of a CL text (Douglas and Hurst 1996 p94).

### **2.2.5 Summary**

Several points of significance for this study should now be noted: many advantages are claimed of CL, but they have yet to be empirically shown; increasing readability is the original motivation for SE, so its rules should be of particular interest; CLs must be acceptable to the reader to be of any use, so surveying whether or not users prefer CL versions is a valuable task.

Now that the field of CL has been introduced, let us look in more detail at the theories surrounding readability.

## ***2.3 Readability***

### **2.3.1 Defining and measuring readability**

Readability is regularly mentioned in the CL literature, but is rarely defined by the authors that use it. Perhaps this is because it is an idea that is prevalent in general language, and authors assume that readers understand the concept as “the ease with which written language can be read with understanding” (Crystal 1992 p326). For this experiment, a more detailed definition was needed to show how readability differed from other metrics, such as legibility, comprehensibility or clarity. At first, it was necessary to consult some fairly dated sources: theoretical work on readability began in the US in the 1940s when literacy levels of the general population were still low, but when the government needed to disseminate increasingly complex written documents in the medical, legal and financial fields. Key works by influential scholars at the time include Rudolf Flesch (1948), Edgar Dale and Jeanne Chall (1948), and Robert Gunning (1952). Dale and Chall (1949 in DuBay 2004 p3) provided the detail lacking in more general definitions. For them, readability is:

The sum total (including all the interactions) of all those elements within a given piece of printed material that affect the success a group of readers have with it. The success is the extent to which they understand it, read it at an optimal speed, and find it interesting.

Klare (1977 in Harkins and Plung 1982 p149) concurs with Dale and Chall in their definition, and states that when we talk of readable writing, “...we mean that the intended readers are able to read it quickly, understand it clearly, and accept it readily (i.e. persevere in reading it)”. In other words, it is the combination of these three elements that differentiates readability from the other metrics. Flesch (1948), Dale and Chall (1948), Gunning (1952), Fry (1968) and Klare (1963) note that we make documents readable to help readers understand them better, and to help them avoid making mistakes that they might otherwise have made. Crucially, though, they emphasise that we also

make them readable to save the readers time and effort, and to ensure that they do not give up on reading the document.

These definitions tell us how to go about measuring readability. However, they do not explain how to predict whether one text is more readable than another. Nor do they instruct us how to produce readable text. These issues are explained in 2.3.2 and 2.3.3.

### **2.3.2 Predicting readability: readability formulas**

As was shown in 2.3.1, measuring what makes a document readable involves the detailed analysis of complex concepts such as the reader's understanding, reading speed, and perseverance. However, such complex analysis may not always be possible. Thus, scholars have tried to develop formulas which use variables in the text to predict how difficult that text may be for a particular audience. Over 50 procedures claiming to compute how difficult a text is to read have been devised over the last 80 years. Of these, six are particularly influential: Flesch; Dale-Chall; Fog; Fry Graph; SMOG; and Automated Readability Index (ARI). Others that are often utilized include FORCAST, Lorge and Spache (Klare 1974 p68).

Formulas do not define or explain readability; they do not point to all the areas of a text that make it readable or comprehensible (Davison and Kantor 1982 pp189-190). The formulas are merely intended as indices or predictors of how difficult a text is likely to be for an intended reader. To construct a formula, the researcher assembles large numbers of 'criterion' passages; these are usually texts taken from the US educational system. Language variables from these passages – typically word difficulty and sentence length – are selected. The researcher then sees how these vary with the scores that readers have given the passages in terms of reading speed, reader preference, and comprehension, to name the three most common values. If a language variable and the readers' scores correlate closely, the variable is said to be a characteristic of readable writing and is combined statistically into a formula. These results are then further validated with other scores for reliability (Klare 1977 in Harkins and Plung 1982 p149).

To use a formula, a passage of at least 100 words is selected; such a length is necessary for the statistical regressions used in most formulas to be valid. Then, a count is made of the language variables that have been identified as being characteristic of readability. These counts are entered into the formula, and an overall score for the passage is given. This score will typically be expressed in different ways: some formulas place the score on a graph (Fry); some express the score as the US grade-school level the reader needs to have completed to be able to read the passage (Flesch); some express the score on a simple scale from 0 to 100, with 0 being the most difficult, and 100 being the easiest (Flesch Reading Ease); while others express it as the number of years of formal education a reader needs to be able to read the passage (Fog).

Crystal (1997 p254) and DuBay (2004 p54) both emphasise the increasing significance and popularity of readability formulas in the field of educational research. However, other authors criticise the formulas for being unsophisticated and unsuited to use on any other texts than those intended for children in the US school system: this is because the criterion passages on which they are based have been selected and validated with schoolchildren in mind (Hargis 2000 p105; Giles and Still 2005 p66).

Now let us look at how readable writing can be produced.

### **2.3.3 Producing readability: the CL created for this experiment**

This section describes the CL created by the author for this experiment: the rules outlined here are shown in the literature to have a positive impact on readability. These rules will be divided into four major categories: textual / pragmatic; syntactic; grammatical; lexical. In her analysis of SE rules, O'Brien (2003 p109) showed that textual considerations outweighed other elements. Remembering from 2.2.3 that readability was the prime motivation behind SE, it is unsurprising that a similar trend appeared in this study.

### Textual rules:

*Have no more than one idea per paragraph.*

According to Davison and Kantor (1982 pp189-191), readability must take into account elements that contribute to a coherent and well-formed text. They emphasise that the inference and cognitive load placed on readers should not be too great.

*Each paragraph should start with a topic sentence.*

Davison and Kantor (1982 pp196-197) also found that reading time shortened and readability increased if closely relevant context information was placed at the beginning of each paragraph.

*Give old information before new (theme-rheme progression).*

Both Farrington (1996 p16) and Reuther (2003 p128) assert that human readers process texts better when new and complex information is presented slowly, in a logical progression, and without too many new chunks at one time.

*Use headings for paragraphs and leave sufficient 'white space'.*

Dayananda (1986 in Crystal 1997 p383) advises writing for the eye as well as the mind: using white space, combined with headings, subheadings, etc., makes the organisation of ideas in the text clearer

*Put long lists in bullet points.*

Hargis (2000 p129) concurs with many CL authors in recommending that long lists should be presented in the form of bullet points.

### Syntactic rules:

In general, Klare (1977 in Harkins and Plung 1982 pp150-151) reminds us that correctly-punctuated 'Simple Active Affirmative Declarative' sentences are the most readable. However, to be more specific, as far as syntax is concerned:

*Sentences should not exceed 25 words.*

O'Brien (2003 p110) explains that in CL the maximum number of words allowed in a sentence varies from somewhere between 20 and 25 words. Generally, the lower limits are applied for procedural texts, and the higher limits for descriptive texts.

*Have variety in the length of sentences within this 25-word limit.*

As was shown in 2.3.1, perseverance is a key pillar of readability. Klare (1977 in Harkins and Plung 1982 pp150-151) underlines that if each sentence is uniformly similar, the reader will become bored and give up reading.

*Have a maximum of two clauses per sentence.*

Bram (1978 in Harkins and Plung 1982 p146) showed that to increase readability, there should only be one or two statements per sentence, with no additional qualifying or explanatory information. In general, difficult texts have a longer more complex structure and impose a greater cognitive load on the reader.

#### Grammatical rules:

*Avoid using ambiguous constructs.*

Some linguistic constructs – for example the connectors 'like' and 'or', or the 'slash' – are ambiguous and require resolution by the reader. These increase reading time and complexity, and should be avoided (Nyberg and Mitamura 1996 p80).

*Avoid using the passive voice.*

Dayananda (1986 in Crystal 1997 p383) states that the passive voice is less readable than the active voice as it generates greater cognitive load.

*Avoid ellipsis and pronominal reference.*

Klare (1977 in Harkins and Plung 1982 pp150-151) states that leaving out parts of sentences and using pronouns - even when the meaning can be understood without the

original noun or ellipited item – creates more difficulty for the reader and should be avoided.

Lexical rules:

According to Nyberg and Mitamura (1996 p77), a pre-approved vocabulary that is consistently used by authors is vital to the success of a CL. However, this predefined word list will vary depending on the domain in which the texts are used: the list for writing a school textbook will be very different to the one used in writing an airplane maintenance manual. No scholar has yet found a better multi-purpose lexical rule for readability than ‘the simple word should be favoured over the complex’. This, however, is not very instructive. Without a vocabulary list specific to this experiment, only one other lexical rule was identified by the author. This was:

*Ensure that all words are spelt correctly.*

Hargis (2000 p129) reminds us that poor spelling can increase processing time, misunderstanding and frustration. Its impact on readability should not be underestimated.

It must be stressed that the above rules should not be accepted without challenge: many are severely criticised in the literature. For example, though bulleting is intuitively held by many to be easier to read, research done by Garrod (1998 in Grover et al. 2000 p91) contests whether doing so actually works. Similarly, Davison and Kantor (1982 pp192-195) claim that shortening sentences can just lead to the dilution of logical relations between clauses and sentences, which in turn leads to mistaken inferences being made by the reader. Moreover, Hargis (2000 p126) asserts that the break-up of sentences not only interferes with understanding in this way, but also produces a choppy, monotonous style that will bore and frustrate the reader. Despite these criticisms, however, the weight of evidence in the literature at present points to the above rules positively impacting on readability.

So far, only the linguistic variables impacting on readability have been dealt with. However, several extra-linguistic variables also have a strong influence on readability and must not be neglected.

### **2.3.4 Extra-linguistic variables**

Many variables outside the linguistic realm help or hinder readers in understanding, in reading more quickly, and in persevering with their reading. These include: motivation; reading ability; interest in the topic; relevance of the topic; familiarity; prior knowledge; and testing conditions. DuBay (2004 p39) points out that many experiments in the field of CL do not achieve the expected results because they fail to control for such variables. It is not difficult to create illustrative examples. Imagine the number of readers that neglect to sign a simple form, even though the instructions to do so are easy to understand and clearly presented: in such a case, it is likely that motivation or interest is lacking. Similarly, we can think of a document that would be completely unreadable or incomprehensible to the average person, but that would be smoothly read and easily comprehended by an expert with prior knowledge of the topic and familiarity with the text type.

Clearly, then, these variables can have an impact on readability. For example, Klare (1977 in Harkins and Plung 1982 p150) shows that:

Someone who is very highly motivated can read very difficult materials, where the mismatch between reading ability and readability is considerable.

He gives some examples: low-ability readers are able to successfully complete a tax return or decipher a complex medical chart when failure to do so would result in serious negative consequences for these readers. Even though motivation and the other extra-linguistic variables are known to be critical, their subjectivity means that most experiments are unable to control for them. As 3.4 of the methodology will show, of the list presented above, this study was only able to examine familiarity and prior knowledge.

### **2.3.5 Summary**

This section introduced the concept of readability and showed that the formulas used to predict readability are much-maligned but are increasingly in use. Moreover, this section explained how readable writing was produced in this experiment, and detailed the extra variables that can impact on readability. It is now time to look at some of the metrics in the field of CL with which readability can be confused.

## ***2.4 Other metrics***

An abundance of metrics in the CL literature has led to conceptual confusion. Some of the main metrics are detailed below.

### **2.4.1 Comprehensibility**

As pointed out by Roturier (2006 p3), comprehensibility should be defined as the ease with which a translation can be understood by its reader. However, in the literature, comprehensibility is often used synonymously with comprehension, understandability and understanding. Is comprehensibility the same as understanding? Few CL works explicitly define the metrics they use in their research and often use terms interchangeably. The conceptual map of the field becomes still more confused when translatability is introduced.

### **2.4.2 Translatability**

This concept is generally taken to be the extent to which a document is amenable to processing by either a human translator or, more often, a machine translation system. However, Reuther (2003) declares readability to be a subset of translatability, while authors like Hargis (2000) reverse this position entirely and see translatability as being just one level of readability.

### **2.4.3 Usability**

Authors like Redish (2000) and Schriver (2000) see usability as a metric which completely excludes the need for readability and comprehension. For them, the aim is to

read to do, or to read to carry out a procedure successfully. They see the reader's level of understanding as unimportant, once the document has been 'used' effectively and the desired result has been achieved.

#### **2.4.4 Others**

Aside from these major concepts, other authors introduce even more 'similar-yet-different' ways to look at CL texts. For example, Puurtinen (1995 p230) defines 'acceptability' as the readability and speakability of a text as well as how well a text receiver accepts a translated text as cohesive, coherent and capable of utilization. She defines 'accessibility' as the ease of comprehension due to the style of writing. Furthermore, Hargis (2000 p123) sees concepts such as 'learnability' and 'doability' as being merely different levels of readability.

Regardless of whether the above definitions are accurate, it is certain that conceptual organisation is required. How might we introduce such terminological rigour to the field?

### ***2.5 Introducing terminological rigour***

The various metrics are highly interrelated: as much as some concepts can appear to contradict each other (as in 2.4), others can be shown to be highly complementary. For example, O'Brien and Roturier (2007) were able to show that many of the CL rules used in their separate studies had a high impact on both comprehensibility and post-editing effort, suggesting that these concepts complement each other.

Thus, a way of mapping CL metrics is required that accounts for such interrelations, complementarities and contrasts. This study proposes the use of a Venn diagram to better understand the conceptual map. Perhaps it will never be possible to draw clear distinctions between what is readable, comprehensible, translatable, etc. Rather than looking to make these concepts entirely distinct, it might be more useful to look at where they have their focus. The metrics in 2.4 focus to a greater or lesser degree on the text

itself, the reader of the text, and the outputs of the text. These three elements then become the three circles of a Venn diagram (see Figure 1).

To illustrate with examples from the diagram: take ‘interest’, Figure 1 illustrates that this metric tells us more about the reader of the text than anything else. In contrast, ‘legibility’ tells us more about the text than anything else. It is by no means the author’s assertion that this is a perfect mapping of the concepts. It is simply intended to convey the opinion that readability can be shown to have a much lesser focus on the reader than comprehension, or that readability can be shown to have a much lesser focus on the results of the text than usability, and so on.

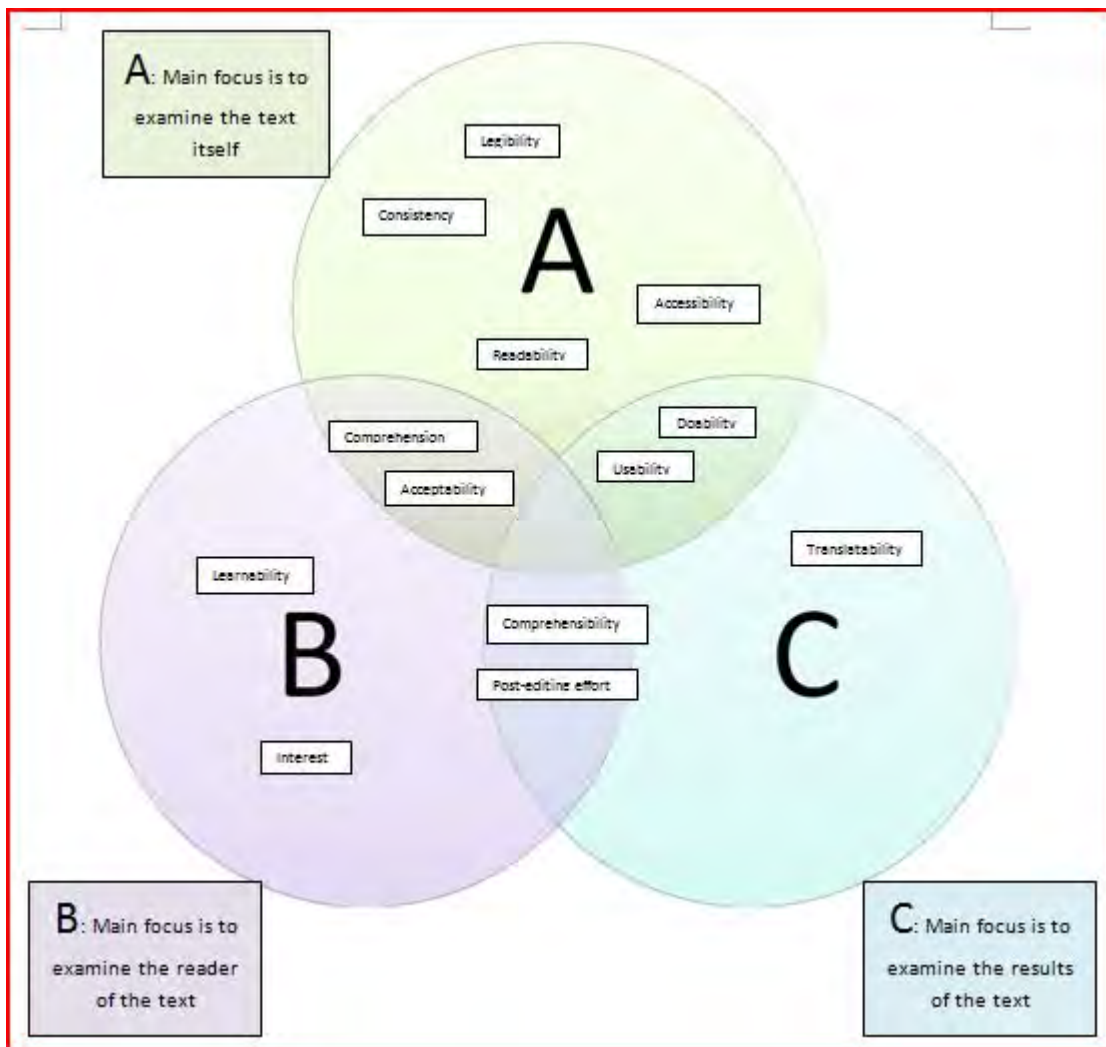


Figure 1: Venn diagram plotting CL metrics

Hopefully, by now, the reader has become convinced that readability is a property of texts that we do well to attempt to control, and that the theory behind it forces us to ask interesting question about text analysis and production. Let us now examine how empirical data were gathered in this experiment to support the ideas presented in this chapter.

## 3.0 Methodology

---

This study began by posing two questions: does readability have merit in the field of CL; and can CL rules be shown to increase readability in texts. It is hoped that the literature reviewed in Chapter 2 has gone some way to answer the first question. To treat the second question, a survey-based experiment was carried out.

### *3.1 Outline of the questions to be answered in the experiment*

The experiment was designed to test whether readers' attitudes to the readability of a sample of technical documents would be influenced by the application of CL rules. This entailed examining various questions:

*Would the CL version be preferred by readers?*

Two versions of the documents were used: one version was a group of extracts from a Symantec manual written in natural language (hereafter NCL version); the other was a version of the same document adapted by the author to conform to the CL rules outlined in 2.3.3 (hereafter CL version). The experiment ran in two stages, with a one-week time-lag between Stages 1 and 2: if a participant was tested on a NCL version in Stage 1, they would receive the CL version of the same passages in Stage 2, and vice versa.

*Would the three pillars of readability be altered in the CL version?*

According to Klare (1977 in Harkins and Plung 1982 p149) the three key elements of readability – efficiency, perseverance and understanding – can be tested by analysing reading speed, reader preference, and retention of key vocabulary respectively. Questions were devised to see if these variables altered from NCL to CL version.

*Would the predictions made by readability formulas correspond to readers' opinions?*

Three passages were selected from the technical manual. These passages were chosen because readability formulas predicted that each one had a widely different level of readability in NCL version and would be more readable in CL version.

*Would extra-linguistic variables impact on readability?*

As was shown in 2.3.4, extra-linguistic variables have a high impact on readability and must be taken into account, where possible. Two groups of people participated in the experiment to see what impact technical expertise, participant profile, and familiarity with the subject matter might have on results. One group was made up of technical staff in Symantec; the other group was made up of participants without such technical expertise and domain familiarity.

*Would certain parts of the texts prove to be more difficult than others?*

Participants were asked not only to give their attitudes, but also to detail what sections in both NCL and CL versions proved hardest to read. This would illustrate whether one version seemed more difficult and whether the areas highlighted corresponded to the areas that did not conform to the CL rules.

Before examining the methodology used to explore these questions in detail, it is important to clarify the theoretical framework upon which this experiment was based.

### ***3.2 Theoretical framework***

The study of attitudes can be carried out using various models. In this experiment, a positivist theoretical framework was adopted. The aim of the positivist is to generate data which hold independently of the research setting and researcher. Such research is about getting statements from informed sources – in this case, through a survey – and comparing the responses: theorists in this framework believe that the more agreement that can be found among participants, the more grounds there is for accepting the statement as true. Equally, it is essential for the positivist to try and secure the stimulus conditions in the experiment through standardisation. It comes out of statistical and behaviourist logic that, by standardizing the research as much as possible, reliable and replicable data are extracted (Silverman 2006 pp119-120). This model assumes that such a thing as a typical respondent exists and that such attitudes and statements can be produced and reproduced.

It is for this reason that the positivist tries to control as many variables as possible in their interventions and attempts to make their observations concrete and inference-free (Silverman, 2006 pp282-283). This positivist theoretical framework conditioned the methodological choices that will now be explained.

### ***3.3 Textual data on which the survey was based***

For this experiment, the author was given access to various naturally-occurring texts in Symantec, “a software publisher specialised in security and availability solutions” (O’Brien and Roturier 2007). To prepare the survey, several tasks needed to be accomplished: select the most appropriate texts; apply them to readability formulas to identify passages with different levels of readability; create a CL version of these passages.

#### ***3.3.1 Selecting appropriate texts to use as NCL versions***

The NCL versions selected by the author came from an internal training manual in Symantec. To give more detail about this document: it contains 24,000 words and is 162 pages long; it was written with both descriptive and procedural functions in mind; it presents a training course designed to be read by the technical staff of Symantec looking for an overview of *Norton Internet Security* products; it is studied in initial training sessions and held for future on-the-job reference; it is written under the assumption that the reader is familiar with domain terminology and wider corporate processes in Symantec; it was written by a team of authors, some of whom do not have English as a native language. Symantec aims for a ‘universal’ style in their document production. However, this manual was only intended for internal consumption and was not as stringently edited as one destined for external clients (personal communication with Johann Roturier 2008). In short, this document seemed ideal for this experiment as it could be claimed that the Symantec participants would all be expert in, or at least familiar with, its contents. Moreover, it seemed long and varied enough to provide different examples of writing styles and readability.

Next, the author divided the manual up into segments. As advised by Flesch (1948 p228), the beginning point for each new passage was aligned with the beginning of a new paragraph. In this way, theme and rheme were more likely to be balanced. A length of 100 to 200 words was chosen for each passage: as per 2.3.2, the regression calculations used in most readability formula require such a length to be valid. Moreover, as it had been agreed with Symantec that the survey would not take more than 20 minutes to complete, reading no more than 600 words would suit such a time frame.

With the document now divided, the author had a supply of passages that were:

- of similar length
- from the same document
- bearing the same function
- with a similar balance of theme and rheme
- in the same domain, with similar terminology
- produced at a similar time, under similar conditions.
- designed to be read together and tailored to various styles of reading

With all the above variables controlled, it was the intention of the author to isolate readability as the main point of difference between the passages. In this way, if attitudes to the passages varied, it could be asserted that such a difference could be attributed to this variable (Silverman 2006 p42).

### ***3.3.2 Isolating different levels of readability***

2.3.1 showed that accurately measuring readability is a complex, time-consuming and subjective task. As this experiment needed to be carried out within a limited time frame, it was deemed appropriate to use popular readability formulas to isolate which passages might have varying levels of readability. In 2.3.2, we saw that six formulas are most influential. From these six, any formulas that base their readability predictions on a crosscheck with a vocabulary list were excluded: these lists are domain-specific – children’s literature, the military, science, etc. – and would be biased against technical

texts such as the ones chosen for this experiment (Hargis 2000 p127). Accordingly, only variations on the Flesch formula (Flesch Reading Ease and Flesch-Kincaid Grade Level) and the Gunning Fog formula remained open as options for this experiment. These formulas were advantageous in that they are in widespread use (Flesch calculators come bundled with *Microsoft Word* applications), they are commonly demanded by government and industry when evaluating documents (Hargis 2000 pp127-128), and they do not require the application of a potentially domain-biased vocabulary list (though they are initially calibrated to standard American grade-school texts, as outlined in 2.3.2).

The author pasted each segment of the manual into a ‘Word’ document. A Flesch Reading Ease and Flesch Kincaid Grade Level score was obtained. Candidate passages were then grouped by score into three main groups: one group that the formulas predicted would be “very challenging” to read; one that would be “somewhat difficult” to read; and one that would be “fairly easy” to read. These candidates were then further analysed by the author to ensure they all had a descriptive, not procedural, function. Promising passages were then entered into an online Fog Index calculator to ensure that this separate index agreed with the Flesch predictions. In the end, three passages were selected to be experimented on. One received an extremely unfavourable readability prediction from all three indices. It will henceforth be referred to as the ‘Norton’ passage. One received an extremely favourable readability prediction from all three indices. It will be referred to as the ‘Shared’ passage. The last passage received a prediction score midway between these two from all the indices and will be referred to as the ‘HijackThis’ passage. The names used here have been adapted from the topic sentence in each passage. The three NCL passages and their readability scores can be seen in Appendix A.

### ***3.3.3 Creating CL versions***

Before proceeding to describe how the CL versions of the passages were created, some important points must be borne in mind by the reader:

- The term ‘CL version’ is used here, but it was not a completely controlled version. The main focus was on applying rules which have been shown in the literature to

impact only readability. Other rules – to improve translatability, comprehensibility, etc. – were mostly ignored.

- Though controlling vocabulary is a vital part of any CL, and complexity of vocabulary can definitely be shown to have an impact on readability, no terminology was controlled by the author in this experiment. This was for two main reasons: to a certain extent, terminological consistency and disambiguation is already guaranteed by the Symantec authoring process, even for an internal manual; and it was necessary to ensure that the content of the passages was still technically accurate after the author's intervention.

To create the CL versions, Symantec gave the author access to acrocheck™, “a quality assurance application with CL checking capabilities” (Roturier 2006 p58). This tool assists writers in conforming to rules concerning spelling, grammar and style. It flags structures that do not comply with a list of formalised CL rules (Roturier 2006 p5). The advantage of working in this environment is that correcting one violation may lead to breaking another rule (e.g. rewriting from passive to active voice may cause a sentence to become too long), and this is immediately flagged by the system. The readability rules applied by the author were detailed in 2.3.3 and are relisted here:

#### Readability rules

*Have no more than one idea per paragraph.*

*Each paragraph should start with a topic sentence.*

*Give old information before new (theme-rheme progression).*

*Use headings for paragraphs and leave sufficient ‘white space’.*

*Put long lists in bullet points.*

*Sentences should not exceed 25 words.*

*Have variety in the length of sentences within this 25-word limit.*

*Have a maximum of two clauses per sentence.*

*Avoid using ambiguous constructs.*

*Avoid using the passive voice.*

*Avoid ellipsis and pronominal reference*

*Ensure that all words are spelt correctly.*

The system also flagged other violations not pertaining to readability, but to other goals, such as increasing comprehensibility, translatability, etc. Specifically, the following other rules were also applied by the author:

Other rules not pertaining to readability also applied

*Avoid using the future tense.*

*Use 'this' or 'that' with a noun.*

*Avoid using 'ing' words.*

*Use 'could' only with 'if'.*

*Keep the verb close to the subject.*

*Use articles.*

*In a conditional sentence, put the 'if' clause first.*

Violations of these rules were corrected by the author because participants would be underlining any areas of the text that they found difficult to read: if it turned out that these areas corresponded to rules not thought to apply to readability, it could make for interesting data.

Once the CL versions had been created, the same readability formulas were applied to them: all three scores improved marginally. However, the relationship between the three passages remained the same (i.e. the most difficult to read remained the most difficult, etc.). Thus, according to the readability formulas, applying CL should make the passages slightly more readable, but not to any great extent. The three CL versions and their readability scores can be seen in Appendix B.

With the NCL and CL passages now ready, it was time to prepare the survey for participants to complete.

### ***3.4 Survey***

The goal of a survey is to enable the researcher to predict accurately the characteristics or attitudes of a predefined group of people. Surveying the entire population is impractical and unnecessary, and a sample of a population can yield highly accurate predictions (Oppenheim 1983 p1). All samples can be expected to misrepresent the population to some degree: the question of the representativeness of my sample will be dealt with in 3.6. Instead, let us first detail how and why the participants were chosen.

#### ***3.4.1 The sample of participants***

This experiment was based on two samples: one sample was made up of participants that had knowledge and practical experience of the technical domain from which the texts were taken (Symantec); the other sample was made up of participants without such domain knowledge and experience (Control). First, the Symantec group was recruited.

##### Symantec

A request was made for 20 volunteers from the technical and language-support sections of Symantec's Dublin office. 12 people responded, participating in both stages of the survey. This was not a random sampling, but a targeted search for people who would be familiar with the type of text under experimentation. Not all were native speakers of English. This would allow the comparison of non-native and native reactions. Other profile data - such as age, education and gender - were also gathered from these participants. With this information in mind, it was then possible to recruit participants for the control group that would mirror the Symantec sample.

##### Control

This mirror sample was put in place to test what impact extra-linguistic variables, like domain expertise or familiarity, might have on results. The author asked 12 associates that fit the necessary profile to volunteer to participate. All did so. Again, this was not a random sample of people, but a targeted search.

As Figures 2 and 3 illustrate, this targeted approach produced two very similar samples. Figure 4 shows that there was some imbalance in the age profiles of the groups. However, ignoring this, it can be asserted that both groups were very similar in all respects, except for familiarity with and expertise in Symantec products and terminology.

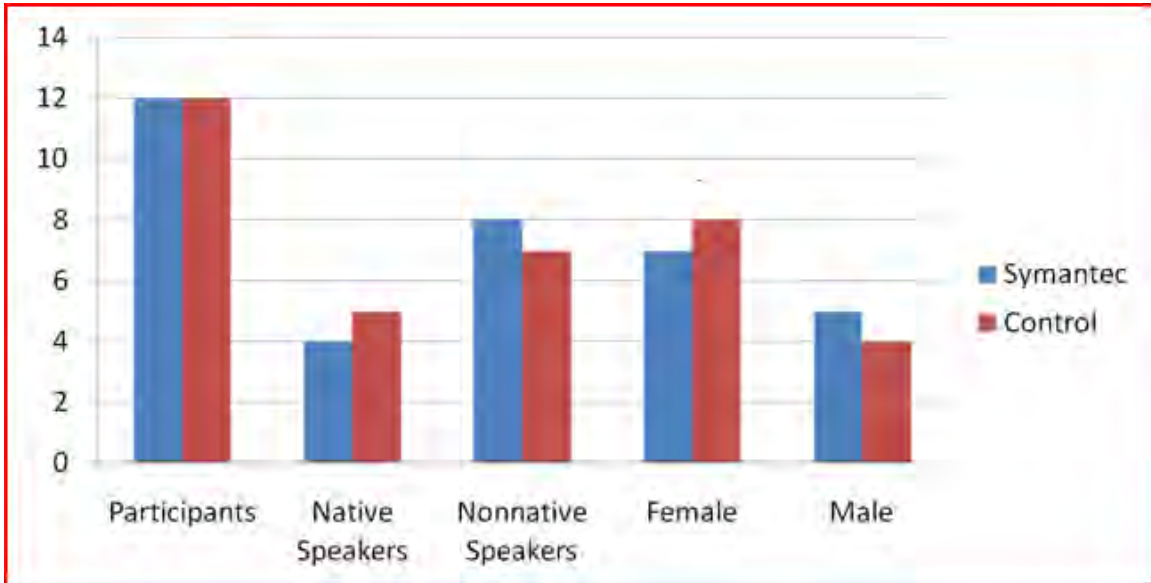


Figure 2: Samples generally similar

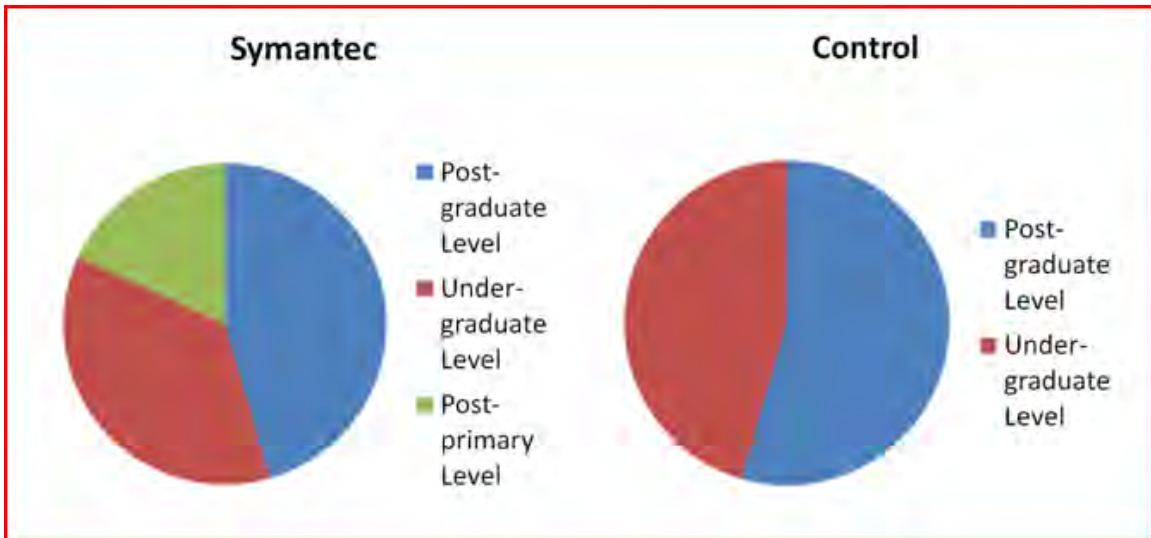


Figure 3: Education fairly similar

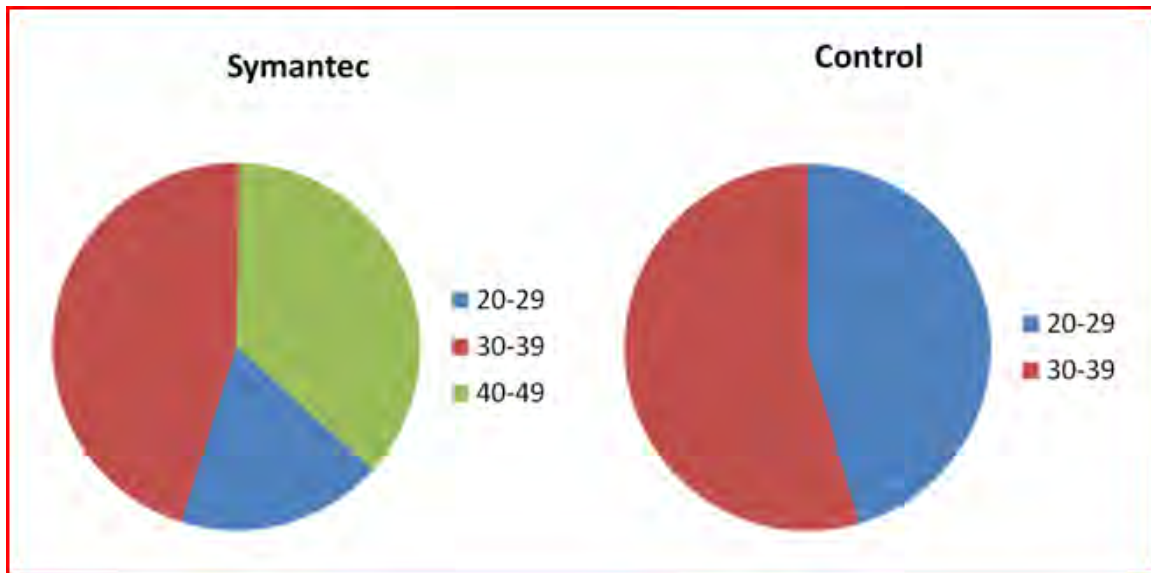


Figure 4: Some imbalance in age

As mentioned in 3.1, this experiment was done on a test / retest basis. There was a survey in Stage 1 and another similar survey in Stage 2 to examine attitude variance depending on whether a CL or NCL version was read. To counter any inherent bias, participants did not know when they received a CL version: for both groups, a table of random numbers was used to randomly distribute a CL version to half the participants in Stage 1 and half in Stage 2. In such a scenario, the time-lag between stages of the experiment to control the memory variable becomes crucial: choosing the length of this time-lag is difficult to justify, as will be explained in the following section.

### ***3.4.2 The time-lag between Stage 1 and Stage 2***

It was hoped that responses at the second reading would be free of the memory of the first reading: i.e., that the first and second-reading stimulus conditions would be identical in all respects except for the CL amendments. To this end, the author also attempted to ensure similar reading conditions at both stages for both groups. However, it would be impossible to expect the participants to have no memory of the first reading experience. The best that could be hoped for was that this impact of memory on the second stage would be minimised. In the literature, it would seem that there are very few psychological theories about forgetting (Wheeler, Ewers and Buonanno 2003 p578). McArdle and Woodcock (1997 p420) assert that a general decline in memory over time is

to be expected, but that there appears to be little agreement as to what variables might impact on this rate of decline. In their study, they expected that memory losses would increase with a longer time-lag between tests. However, this hypothesis was not substantiated. In addition to time, other variables have a great impact on the rate of memory loss. Wheeler, Ewers and Buonanno (2003) agree with these findings and claim that other variables such as encoding conditions and the need or likelihood of re-use vastly impact on the rate of forgetting. Similarly, Cowan (2000) declares that the number of chunks of information to be retained is central to the rate at which memory loss will occur. Yet, there is little agreement in the literature as to what might be the mathematical rate of forgetting, and even less about what variables contribute to this decay. In all the experiments noted above, the time periods examined ranged from several minutes to several days, with a general trend (though not an uninterrupted one) of decreasing memory as time-lag increased. An interval of 48 hours between test and retest is considered long and, thus, in this present experiment a time-lag of 7 days seemed appropriate.

Now that we have seen how the survey was set up, let us look in detail at what questions were asked of participants and why.

### ***3.5 The questionnaire***

Several general principles guided the design of this questionnaire. Specific rather than general questions were given precedence. Converse and Presser (1986 pp31-3) state that such a format tends to encourage better recall, and it has been shown to be a better predictor of the behaviour of the variables being studied. Questions relating to actual rather than hypothetical situations were prioritised, as is recommended for the positivist model (Silverman 2006 p119). Closed questions were favoured over open ones, in order to communicate the same frame of reference to all participants (Converse and Presser 1986 p33). As Silverman (2006 p119) asserts, it is essential in qualitative surveys to check first whether participants have beliefs about the topic beforehand to avoid putting words in their mouths: thus, the questionnaire was designed to proceed from attitudinal questions about readability to more specific questions about the texts. Finally,

transitional statements guiding participants through the questionnaire were clearly presented (Oppenheim 1983 p67).

In 2.3 and 2.4, it was shown that, though highly interrelated, readability is not exactly the same as comprehensibility, translatability, etc. This point was emphasised by the author in the preamble to the questionnaire. The phrase ‘To be or not be’ was given as an example: this sentence is extremely readable (it is read speedily, with high acceptance, and with high retention). However, can we say that this sentence is easy to comprehend? A school student, a professor of English and a Shakespearean actor may all understand this phrase in very different ways. Moreover, would this be easy to translate into another language? And would all translators do it in the same way? Participants were requested to remember at all times that the questionnaire was asking them how easy (or not) certain passages were to read.

Let us move now to examine each individual question. A sample questionnaire is included in Appendix C.

### Question 1

*Please read the following statement “It is important for training documents to be easy to read.” and put a tick under the option which best describes your attitude (Strongly Agree; Agree; Disagree; Strongly Disagree; No opinion).  
If you agreed or disagreed with this statement, why do you feel this way?*

Oppenheim (1983 p37) emphasizes that first questions must grab respondents' attention and help motivate them to continue. Also, they are a way of getting unbiased general attitudes before the more specific goals of the questionnaire become clear to the participants. Thus, the questionnaire began by placing the focus immediately on a central issue of the study: whether or not readability is important. To standardise the stimulus conditions and to make sure that the participants were answering the questions as intended by the author, it was necessary to define what was meant by training documents in this case. Furthermore, by making this a closed followed by open-ended “why” question, the standardization and filtering advantages of a closed format, along with the

potential for rich data of an open format could both be accrued (Converse and Presser 1986 p43).

### Question 2

*Please tick the option which best describes the way you read training documents (Skim; Scan; In-depth). If you want to choose more than one option, please give more detail.*

The way that a person will read a text has great relevance to questions of word and sentence length, sentence and paragraph organization, information load, etc. These variables often correlate with readability. This question was asked to see if any significant trend was noticeable in the way that readers usually read training documents.

### Question 3

*Please match each concept in Column A to the most appropriate definition in Column B (Comprehensibility, readability and acceptability were then defined). Draw a line from the concept to the definition to indicate your choice.*

Though it was emphasised in the briefing session and in the instructions that this was an experiment into readability, it was feared that many participants would still have ideas of comprehensibility or other similar concepts in mind when answering the questions. This question was inserted to reemphasise to the participant that readability is different to other metrics.

### Question 4

*Please select the option which best describes the passage you have just read: It was fairly easy to read; It was somewhat difficult to read; It was very challenging to read.*

Participants were asked at this stage to read the three passages and rate them for readability. Unfortunately, at this stage the author neglected to include a “No opinion” option (see 3.7 for more discussion). This question was to show whether participants agreed with the predictions made by the readability formulas and whether there was consistency among participants. It was hoped that it would also show whether estimations of readability were different depending on whether CL was used or not. Another vital piece of data gathered at this stage was an assessment of the amount of time

it took each participant to read each passage. As emphasised in 2.3.1, reading speed is a key pillar of readability. Each participant simply noted their beginning and ending times for reading each passage. It is important to stress here that the passages were not presented in the order ‘most to least difficult’. This was to control for the idea that once the most difficult text was completed, the others might seem easier by comparison. Or vice versa, had the easiest been presented first, the others might have seemed relatively more difficult. Rather they were presented ‘Hijack’ (middle-ranging readability prediction) to ‘Norton’ (most difficult readability prediction) and ‘Shared’ (least difficult readability prediction).

#### Question 5

Here the same passages were provided a second time. On this occasion, participants were asked to underline parts of the text (if any) that they found difficult to read. This task was made into a separate question at a second reading to avoid cognitive overload. This question was an attempt to see if the areas that violated various CL rules corresponded to areas that participants found difficult to read.

#### Question 6

*You have now read three passages about Norton Internet Security 2007. Have you ever read these passages before today? (Yes; No; Don't remember)*

This question was only asked of Symantec Group and was to control for the fact that the author might have unintentionally chosen a passage that they use in their jobs every day. Such a level of familiarity (another extra-linguistic variable) would have a huge impact on attitudes to readability.

#### Question 6b

*(Asked of both groups) Did any of the following words appear in the three passages? If so, please tick as appropriate. (Safety, Shares, Tool, Device, Security, Allocation)*

As discussed in 3.1, retention in short-term memory of keywords and concepts has been identified as a way of categorising understanding in readability: if key items are retained easily, the passage is said to have been ‘understood’ better and is more readable. Using *Wordsmith* corpus tools, three keywords (one from each passage) were selected: Security;

Shares; Tool. Three synonyms that did not appear in any of the passages were then added to the list: Safety; Device; Allocation. All words on the list were the same part of speech (noun). All were from the same technical domain found by looking in a technical language dictionary. If a synonym appeared in the text in another part of speech it was discounted. The keywords remained the same for the CL versions of the text, and it was ensured that the synonyms had not been introduced during the CL text creation. In the instructions, participants were requested not to turn back to the previous page to check their answers.

As such a list prompts readers without any other grammatical or syntactic hints, it can be said to be a pure test of retention. This question was not testing participants' comprehension of the key concepts or words: it was largely a pattern-matching exercise. To retain this item in short-term memory, it is likely the participants had some understanding of a concept like 'security' or 'share'. This is not to say that they necessarily understood these concepts as used in the computer domain, or more particularly in Symantec. It must be recognised that this is not a perfect test: by now in the questionnaire, participants had read the texts twice; moreover the last text appearing before this question probably had a retention advantage.

#### Question 7

*Please tick the most appropriate choice: I think the version from the last time is easier to read; I think the version from today is easier to read; I don't know.*

This last question appeared only in Stage 2. It included the essential 'No preference' option and was used to see whether readers preferred the CL version. If this was the case, it could be said that CL increased readability.

It is now time to make clear the various limitations that must be recognised in this methodology and these data.

### ***3.6 Limitations***

The participants in this experiment represent a very small, non-random sample. Thus, it is difficult to generalize any results to a larger population of users of CL texts. This study can only claim to be a pilot study to indicate future fruitful avenues of research.

To give their informed consent, all participants had to be aware that this was an experiment into readability and CL. People working in the technical domain can hold significant bias – both positive and negative – toward CL. Though this experiment attempted to control for such bias, as explained in 3.1, the effect of participants’ preconceived ideas about the topic cannot be discounted from any analysis.

This methodology depended on participants reporting truthfully, consistently and accurately. However, Oppenheim (1983 pp105-109) reminds us that when dealing with attitudes, these are the very factors that are most difficult to guarantee. Moreover, as the reading of passages was self-timed, these questions were particularly open to inaccuracy. Finally, this experiment really only sets out to establish that some relationship can be shown to exist between CL and readability; with a methodology based on survey data, we cannot say that this relationship is undeniably one of causality (Oppenheim 1983 p6).

Having now recognised that the data and methodology in this experiment are limited, it is possible to go further and recommend methodological steps to be avoided in future experiments.

### ***3.7 Methodological points to avoid***

Converse and Presser (1986 p36) warn that offering three choices for an answer leads to non-committal responses and should be avoided. Unfortunately, in Question 4 of the survey, participants were presented with such a choice and it proved highly ineffective. Worse still, the option ‘very challenging’ was probably too extreme a classifier for this question: people may have been reluctant to choose something so strong and loaded with negative connotations. Overall, the idea of using adjectival classifiers proved to be very

unrefined. Such limited choices provided no way of tracking changes in attitudes: for example, a ‘fairly easy’ passage was still a ‘fairly easy’ passage at the second stage, though this may have been hiding the fact that the participant actually found it significantly easier. In future studies, a numbered scale of difficulty – perhaps rating passages from 1 to 10 – would surely prove to be a more effective tool.

Equally, the method of timing used when participants were reading was crude and probably inaccurate. Asking people to note starting and finishing times might be sufficient for very long time periods, but in this experiment, with such short passages, accurate counting of seconds and even milliseconds would have provided much richer data. In this way, the use of eye-tracking software, with its accurate time measurement and complex reading-pattern display, would bring great benefits.

Finally, any future study would do well to reconsider the readability formula used. This author chose formulas that are the most well-known, the most widely available, and the most easy-to-use. As the experiment progressed, the many weaknesses of these formulas became clear: they do not control for vocabulary in their predictions; they are calibrated to school textbooks; and they are correlated to scores given mainly by school children and teachers in the US educational system. In very recent works, both Klare (2000) and DuBay (2004) recommend The New Dale-Chall Readability Formula (see Chall and Dale 1995). It is calibrated to school texts *and* technical papers, it is tested on scores given by children *and* adults, and it requires passages being studied to be rated for difficulty against an extensive vocabulary list. Therefore, it would seem that this formula may be able to give more accurate predictions than those achieved here.

By now, the reader should understand why and how the data in this experiment were gathered. In the next chapter, we will examine whether these data answered the questions posed at the outset.

## 4.0 Data Analysis

---

There are many ways to present the data analysed here: the themes that data have been grouped under correspond to the research questions that were laid out in 3.1. A tabulation of all the raw data collected in this study can be consulted in Appendix D. Let us look, now, to the first question.

### 4.1 *Would the CL version be preferred by readers?*

At Stage 2, after having read both CL and NCL versions of the text, participants were asked which one they found easier to read. A majority of participants (see Figure 5) in both groups said that they found the CL version of the texts easier to read. To this extent, it can be claimed that the CL versions were preferred by these readers.

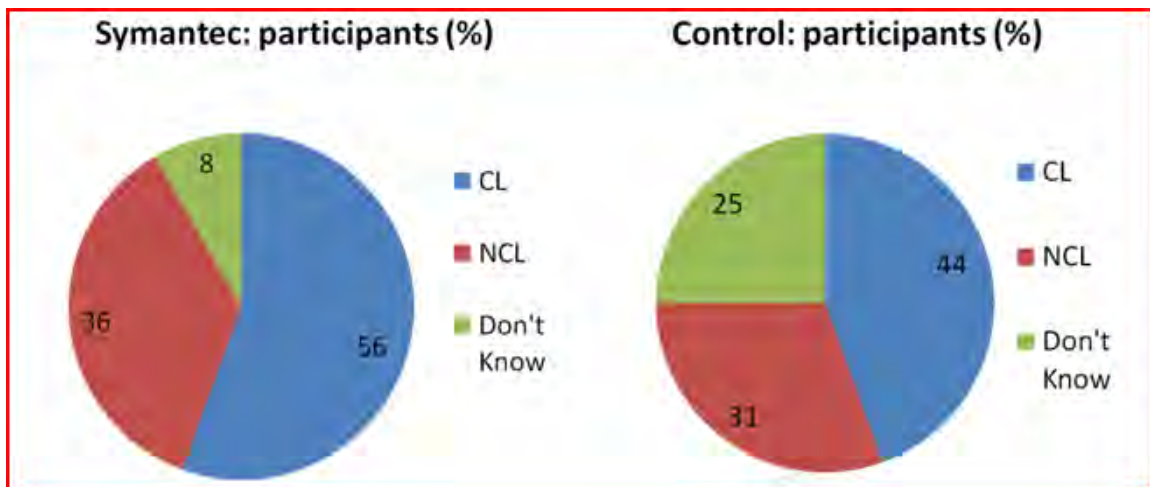


Figure 5: Which texts were easier to read – what % of participants chose what

To further test for preference, a value was assigned to each classifier that participants used to rate the texts. If a participant selected 'Fairly Easy' to rate a text, this was given a value of 1. 'Somewhat Difficult' was given a value of 2. And 'Very Challenging' was

given a value of 3. These totals were then added up and CL was compared to NCL as in Figure 6.

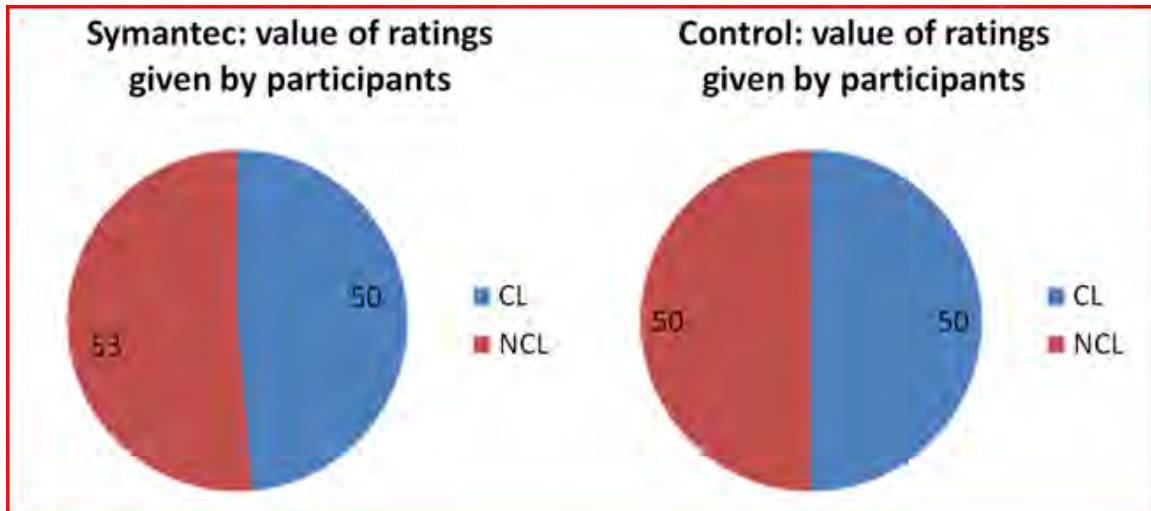


Figure 6: Favourability of ratings - CL vs. NCL

Unfortunately, these data are inconclusive: this is likely a function of the inappropriate methodological choice explained in 3.7. A more refined rating scale would have possibly shown more significant favourability differences between CL and NCL versions.

Despite these inconclusive data, the evidence in Figure 5 suggests that overall the CL version was preferred by readers. If a version is preferred by a reader, they are more likely to persevere with reading it than another version. So, on this level, we can say that CL versions appear to be more readable. But what do the questionnaires tell us about the other two pillars of readability?

#### ***4.2 Would the other two pillars of readability be altered in the CL version?***

The first point we will examine is the retention of key vocabulary. After reading all the passages, participants were asked to identify keywords from a list including synonyms that did not appear in the passages. Figure 7 breaks down the number of correctly retained keywords - less incorrectly selected synonyms (noise) – for each group at each stage.

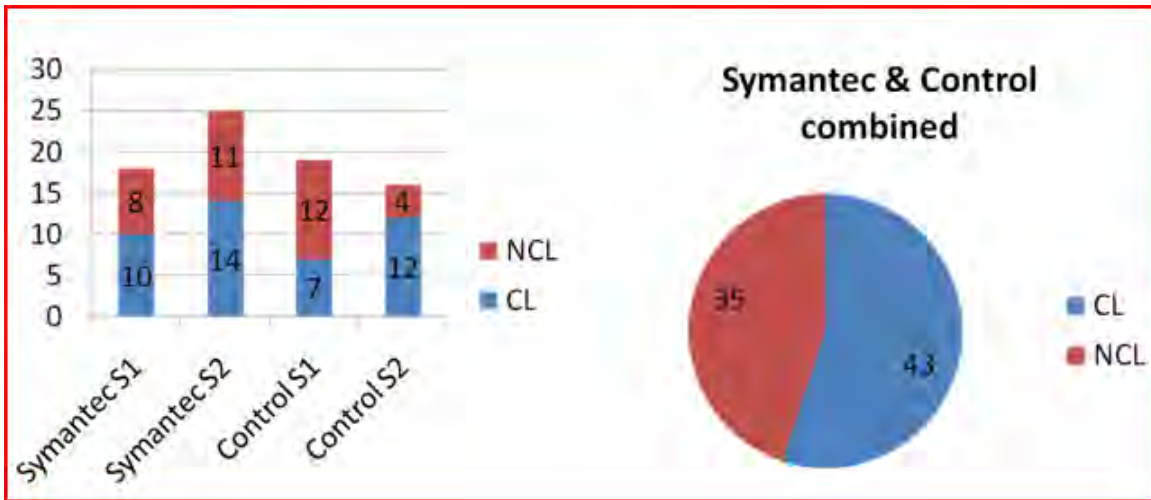


Figure 7: Number of keywords correctly retained minus noise

Overall, this shows that more key vocabulary was correctly retained with less noise when the CL version was read. This result indicates that retention is better when a CL version is used, and that, on this level, the CL texts were more readable. However, as we know from 2.3.1, readability is about more than just retention and perseverance. The last element to be tested was reading speed.

In fact, Figures 8 and 9 show that in the majority of texts, for both groups and at both stages, the CL versions actually took longer to read. This was an unexpected result. However, as was noted in 3.7, due to the crude and likely inaccurate self-timing, these speed figures are not to be relied on too heavily. It would be interesting to see the results a more accurate timing methodology would give.

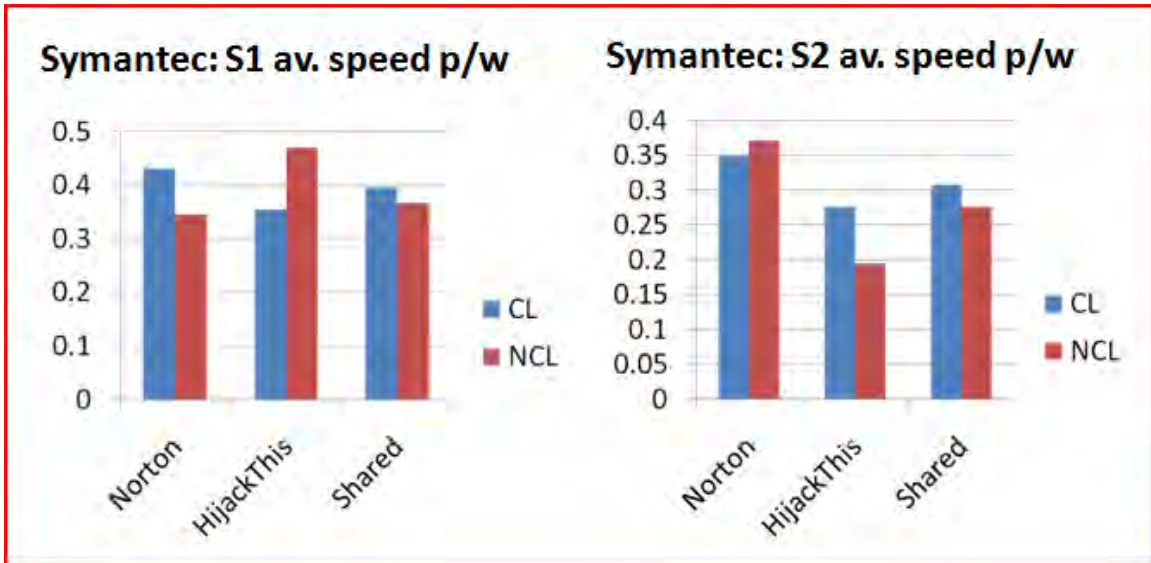


Figure 8: Breakdown of average reading speed per word per text for Symantec (seconds)

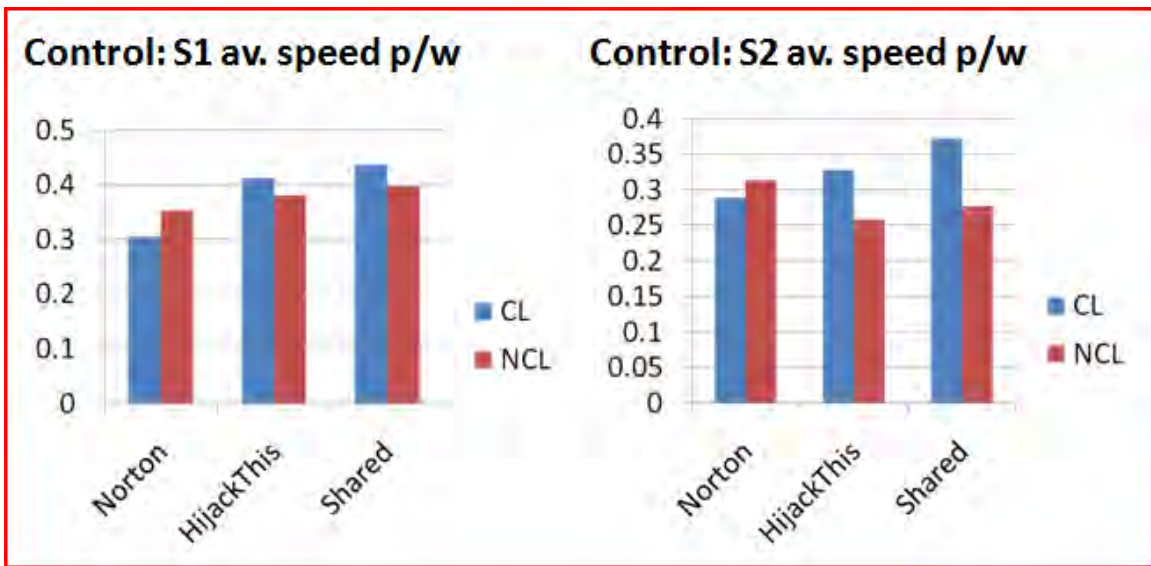


Figure 9: Breakdown of average reading speed per word per text for Control (seconds)

In summary, speed was not positively impacted by applying CL rules, though retention and perseverance were: therefore, it would seem that applying CL rules to technical documents does increase readability.

### 4.3 *Would the formulas' predictions correspond to readers' opinions?*

Overall, it seemed that the predictions made by the formulas did not correspond to what readers thought. Figure 10 illustrates that for both groups the majority of readers' ratings did not correctly correspond to the readability formula predictions.

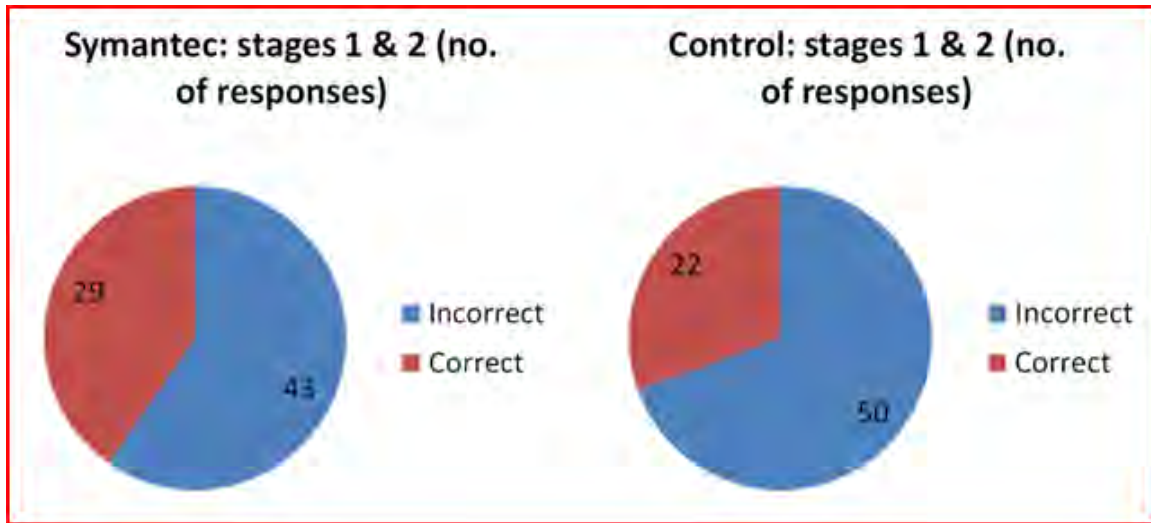


Figure 10: No of times readers' ratings correctly corresponded to formula predictions

The formulas chosen do not appear to be useful predictive tools. However, there was another group of data that told a different story: participants were asked in the questionnaire to underline any parts of the text that they found difficult to read. By very crudely counting the number of texts in which participants underlined something for reading difficulty, it can be shown that the formulas actually predicted the relationship between the three passages correctly. Thus, in Table 1 we see that in the naturally-occurring NCL versions, 'Norton' had the most responses with underlined sections (17), 'Shared' had the least (11), and 'HijackThis' came in the middle (13). This corresponds exactly to the ranking for difficulty that the formulas predicted in 3.3.2.

Total no. of texts with underlining for difficulty				
		Symantec	Control	TOTAL
NCL	Norton	11	6	<b>17</b>
	HijackThis	6	7	<b>13</b>
	Shared	5	6	<b>11</b>
CL	Norton	6	5	<b>11</b>
	HijackThis	6	7	<b>13</b>
	Shared	6	7	<b>13</b>

Table 1: No. of texts with underlining for difficulty

Of course there is a ‘black-box’ problem with the validity of these data. That is, the underlined sections can only be said to represent areas that participants were less satisfied with; we do not know whether people were truly underlining for problems of readability, comprehension, or some other objection (type-face, legibility, subject matter, etc.). Despite this validity issue, the underlined sections still raise important issues about the way of testing reading difficulty: asking an opinion or rating is subjective and value laden. Perhaps task-related testing, such as the underlining task, generates more objective data.

#### ***4.4 Would extra-linguistic variables impact on readability?***

In 2.3.4, prior knowledge of a domain (or technical expertise) was identified as an extra-linguistic variable that is shown to increase readability in the minds of readers. By this hypothesis, then, the Symantec group should have given more favourable ratings than the Control group. However, this was not the case. Using the same methodology as in 4.1, the Control’s total favourability score was 100 to Symantec’s 103. Again, these data suffer from the lack of sophistication of the rating system chosen. Nonetheless, Table 2 shows that Symantec underlined more for difficulty than Control, again pointing to the fact that prior knowledge of the domain was not positively impacting on readability. It must be stressed, though, that no Symantec participant answered in Question 6 that they

remembered reading the texts before. This calls into question how detailed the participants prior knowledge actually was.

<b>Symantec: breakdown of underlining (no. of texts)</b>					<b>Control: breakdown of underlining (no. of texts)</b>						
	Stage 1		Stage 2		TOTAL		Stage 1		Stage 2		TOTAL
NCL	Norton	6	Norton	5	<b>11</b>	NCL	Norton	4	Norton	2	<b>6</b>
	HijackThis	4	HijackThis	2	<b>6</b>		HijackThis	4	HijackThis	3	<b>7</b>
	Shared	4	Shared	1	<b>5</b>		Shared	3	Shared	3	<b>6</b>
					<b>22</b>						<b>19</b>
CL	Norton	1	Norton	5	<b>6</b>	CL	Norton	2	Norton	3	<b>5</b>
	HijackThis	4	HijackThis	2	<b>6</b>		HijackThis	5	HijackThis	2	<b>7</b>
	Shared	3	Shared	3	<b>6</b>		Shared	3	Shared	4	<b>7</b>
					<b>18</b>						<b>19</b>

Table 2: No. of texts underlined for difficulty - Symantec vs. Control

It has been shown here that prior knowledge did not seem to have a positive impact on readability. In contrast, the opposite effect was shown for familiarity. Table 3 illustrates that, in both groups, from Stage 1 to Stage 2 underlining decreased, favourability improved, and speed decreased. This would seem to suggest that – because it occurred equally in Control and Symantec – just becoming familiar with a text, even if you do not necessarily comprehend it or use it effectively, makes that text seem more readable to you.

Familiarity had an impact on readability					
		Symantec		Control	
		Stage 1	Stage 2	Stage 1	Stage 2
Underlining decreased		22	18	21	17
Favourability scores improved		54	49	53	47
Reading speed improved	CL	0.39395	0.31128	0.38513	0.33023
	NCL	0.39483	0.28092	0.37756	0.28336

Table 3: Familiarity appeared to have a strongly positive impact on readability

The final extra-linguistic variable that was tested by this study was participant profile. In particular, native-English ability was shown to have a strong impact on views of readability. By comparing Figures 11 and 12, it can be seen that non-native speakers found the CL versions easier to read, while native speakers tended to find the NCL versions easier to read. This is not a criticism of the non-native speakers' English ability: most rated themselves 8, 9 or 10, where 10 represented native-level fluency. This difference most likely comes about because native speakers are much more familiar with, and tolerant of, the eccentricities and exceptions of naturally-occurring language. Such a finding could also have implications for a company like Symantec, with a multi-national customer base, some of whom may not be reading in their native tongue.

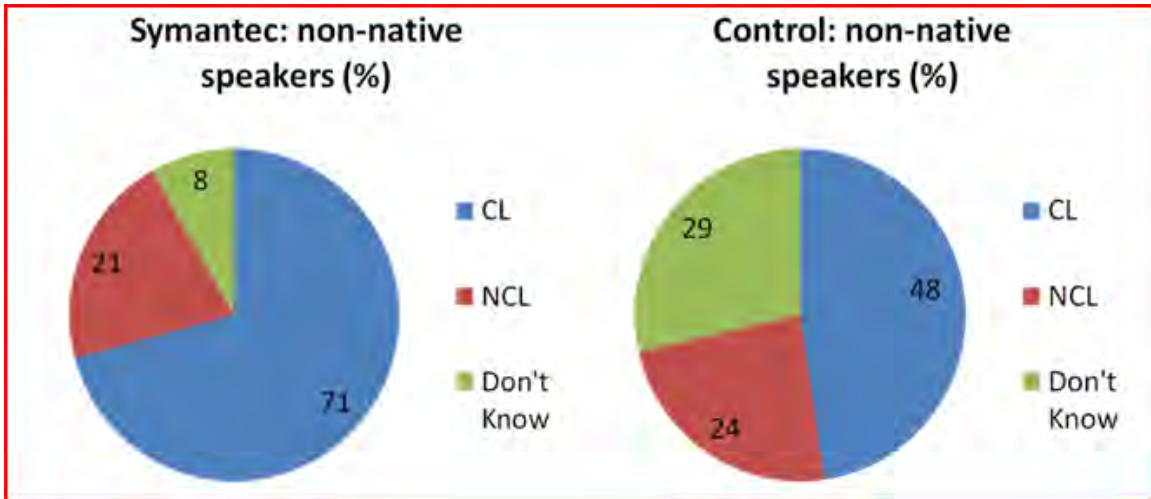


Figure 11: Non-native speakers found CL more readable

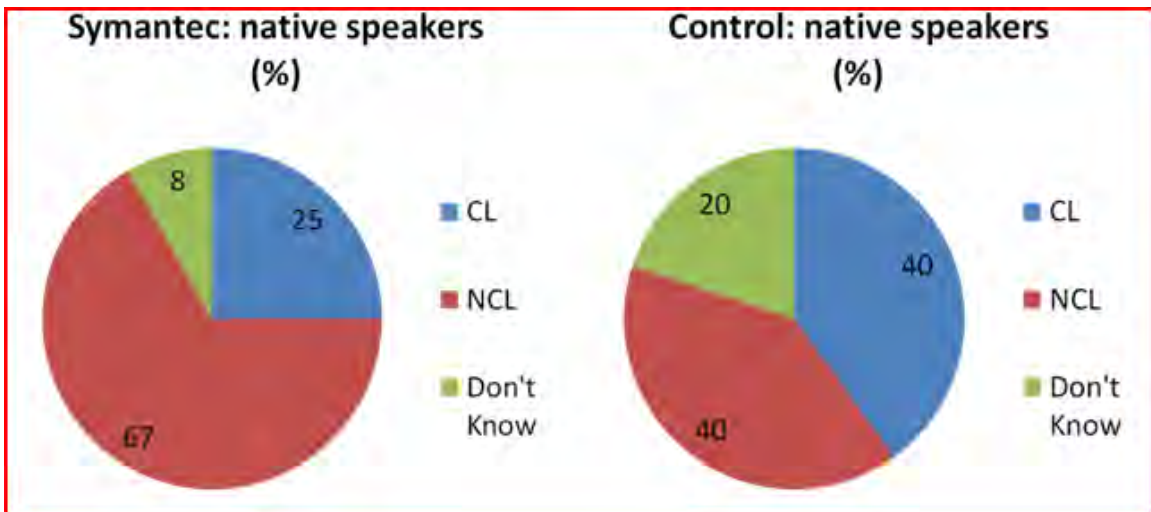


Figure 12: Native speakers found NCL more readable

#### ***4.5 Would certain parts of the texts prove to be more difficult than others?***

It was hoped in this section to match up the underlined areas of difficulty to the sections of the text that had been controlled for readability. However, a comparison of all the underlined data did not reveal any noticeable trends. Moreover, it was decided by the author not to continue with this analysis because of the ‘black-box’ problem detailed in 4.3. Nonetheless, the underlined data are available for consultation in jpeg form at: [http://s290.photobucket.com/albums/II252/Pakkun\\_photo/Symantec%20and%20Control%20Underlining/](http://s290.photobucket.com/albums/II252/Pakkun_photo/Symantec%20and%20Control%20Underlining/)

#### 4.6 Surprising extra data

One unexpected finding came out of this experiment. When retention of key vocabulary was tested, noise was introduced by the author. The overall level of noise (number of times that a synonym that never appeared in any text was identified by the reader as having been in the text) was very constant throughout both stages and in both groups. And yet, one piece of noise was selected disproportionately (see Table 4).

Breakdown of noise				
		Stage 1	Stage 2	TOTAL
Symantec	Safety	5	5	<b>10</b>
	Device	2	2	<b>4</b>
	Allocation	1	1	<b>2</b>
Control	Safety	7	7	<b>14</b>
	Device	4	2	<b>6</b>
	Allocation	0	3	<b>3</b>

Table 4: Breakdown of noise when retention was tested

That ‘safety’ should be chosen so many times indicates an interesting terminological point: people clearly associate *Norton Internet Security* with the word ‘safety’. Such an occurrence may seem natural in the Control group made up of the general public; however, Symantec defines itself as specialising in ‘security’ solutions, so it was surprising that Symantec employees should consistently identify ‘safety’ as a keyword.

In addition, this experiment unintentionally proved that a one-week lag between test and retest is insufficient for having a fresh reading experience in the second stage. The familiarity variable was shown to have a strong impact on results; any study wishing to cancel out this effect would need to introduce a considerably longer time-lag between tests to be effective.

## 5.0 Conclusion

---

It has been shown in this paper that the study of readability is a useful exercise in the field of CL. By making us consider key elements like reading speed, reader perseverance and reader understanding, as well as influential external factors like motivation and familiarity, it promotes a comprehensive approach to the theory of document production. Moreover, this study has provided new empirical data, albeit limited in scope, to show that CL versions: are thought to be easier to read; are viewed more favourably; and encourage better retention of keywords. In short, these data seem to suggest that the application of CL rules increases readability. However, these are not the only conclusions that should be drawn from this experiment.

One such additional conclusion concerns readability formulas. Overall, this study appeared to show that the formulas made inaccurate predictions. However, the underlining task in the survey then cast doubt on this by suggesting that the formulas might predict which texts users would be the least satisfied with. This indicates that it would be a mistake to completely discount readability formulas as useless. As noted in 3.7, the New Dale-Chall formula could be an interesting prospect for future research. Most criticisms of formulas probably arise because of people incorrectly using them in ways for which they were not intended: as guidelines for writing or as tools for correction. If the formulas are used within their recognised limitations – as rough predictors of difficulty to prompt more detailed textual analysis – then they may have more merit.

An important conclusion that needs to be made, too, is that much work remains to be done on clarifying the many metrics that are used in the field of CL. The literature abounds with such concepts as readability, translatability, comprehensibility, usability, etc. that are often confused and misused. The ideas presented in 2.5 are a tentative first step at disentangling the web. Rather than thinking that one form of analysis is better or worse, the suggested approach should encourage people just to consider the appropriateness of each metric to their situation. To give an example, let us ask what

form of analysis might be appropriate for Symantec: this company produces vast amounts of documentation, in short time-frames, for a highly diverse readership who generally just want to solve problems. As such, it is unlikely they would want to use a metric that tells them a lot about their readers, or how they comprehend or learn from their texts. Rather Symantec would do better to focus on metrics that tell them about their texts more subjectively: metrics that focus on the outputs of their texts - on what readers actually do with the texts - would be the most fruitful. Thus, the analysis presented in 2.5 would suggest that 'usability' or 'doability' might be the best way forward. Of course, this is just one example, and the terminological conclusions drawn in this paper are by no means final: it is the author's hope that future studies will focus on more clearly delimiting the conceptual field in the area of CL.

The final conclusion concerns other recommendations for how future studies might proceed. Clearly, more polarised texts will produce more noticeable trends in the data: the documents presented in this test were selected only based on the predictions of readability formulas. Future studies should incorporate not only formulas, but also semantic assessments, expert advice from the users or authors of the texts, and other criteria to decide the difficulty of the passages to be experimented on. Similarly, such tests should try to incorporate extra-linguistic variables into their methodologies. This is a challenging task that requires careful consideration. Perhaps by choosing participants with different levels of responsibility in a company, or by selecting a mix of new-hires and experienced employees, ideas of motivation, familiarity and relevance could be accounted for.

It is hoped that, by now, the reader agrees that readability is a somewhat misunderstood and undervalued concept. This experiment intended to show that controlled-language rules improve readability; but it is further hoped that it has shown that the theory of readability can improve the field of controlled language and the document production process.

## Reference List

---

Chall, J.S. and Dale, E. 1995. *Readability revisited : the new Dale-Chall readability formula*. Cambridge, Mass.: Brookline Books.

Converse, J.M. and Presser, S. 1986. *Survey questions: handcrafting the standardized questionnaire*. Beverly Hills, California: Sage.

Cowan, N. 2000. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and brain sciences*. 24, pp87–185.

Crystal, D. 1992. *An encyclopedic dictionary of language and languages*. Oxford, UK; Cambridge, Mass., USA: Blackwell.

Crystal, D. 1997. *The Cambridge encyclopedia of language*. 2<sup>nd</sup> ed. Cambridge; New York: Cambridge University Press.

Dale, E. and Chall, J.S. 1948. A formula for predicting readability: Instructions. *Educational research bulletin*. 27 (2). pp 37-54.

Davison, A. and Kantor, R.N. 1982. On the failure of readability formulas to define readable texts. *Reading research quarterly*. 17 (2), pp187-209.

Douglas, S. and Hurst, M. 1996. Controlled language support for Perkins Approved Clear English. *IN: Proceedings of the first controlled language application workshop*, March 1996. Leuven: Centre for Computational Linguistics, pp93-105.

DuBay, W.H. 2004. The principles of readability [Online]. Available from: <http://www.impact-information.com/impactinfo/readability02.pdf> [Accessed 2 September 2008].

Farrington, G. 1996. AECMA Simplified English: an overview of the international aircraft maintenance language. *IN: Proceedings of the first controlled language application workshop*, March 1996. Leuven: Centre for Computational Linguistics, pp1-21.

Flesch, R. 1948. A new readability yardstick. *Journal of applied psychology*. 32 (3), pp221-233.

Fry, E.B. 1968. A readability formula that saves time. *Journal of reading*. 11, pp513-516 and pp575-578.

Giles, T.D. and Still, B. 2005. A syntactic approach to readability. *Journal of technical writing and communication*. 35 (1), pp47-70.

Grover, C., Holt, A., Klein, E. and Moens, M. 2000. Designing a controlled language for interactive model checking. *IN: Proceedings of the third international workshop on controlled language applications*. April 2000. Seattle: ANLP/NACLA, pp90-104.

Gunning, R. 1952. *The technique of clear writing*. New York: McGraw-Hill.

Hargis, G. 2000. Readability and computer documentation. *ACM journal of computer documentation*. 24 (3), pp122-131.

Harkins, C. and Plung, D.L. 1982. *A guide for writing better technical papers*. New York: IEEE Press.

Hayes, P., Maxwell, S. and Schmandt, L. 1996. Controlled English advantages for translated and original English documents. *IN: Proceedings of the first controlled language application workshop*, March 1996. Leuven: Centre for Computational Linguistics, pp84-92.

Huijsen, W. 1998. Controlled language: an introduction. *IN: Proceedings of the second international workshop on controlled language applications*. May 1998. Pennsylvania: Carnegie Mellon University, pp1-15.

Klare, G.R. 1963. *The measurement of readability*. Ames, Iowa: Iowa State University Press.

Klare, G.R. 1974. Assessing readability. *Reading research quarterly*. 10 (1), pp62-102.

Klare, G.R. 2000. Readable computer documentation. *ACM journal of computer documentation*. 24 (3), pp148-168.

Knops, U. 2000. Efficient roll-in and roll-out of controlled language applications. *IN: Proceedings of the third international workshop on controlled language applications*. April 2000. Seattle: ANLP/NACLA, pp134-135.

McArdle, J.J. and Woodcock, R.W. 1997. Expanding test-retest designs to include developmental time-lag components. *Psychological methods*. 2 (4), pp403-435.

- Means, L., Chapman, P. and Liu, A. 2000. Training for controlled language processes. *IN: Proceedings of the third international workshop on controlled language applications*. April 2000. Seattle: ANLP/NACLA, pp1-13.
- Mitamura, T., Baker, K., Nyberg, E. And Svoboda, D. 2003. Diagnostics for interactive controlled language checking. *IN: European association for machine translation: international workshop*, May 2003. Dublin: Dublin City University, pp87-94.
- Moller, M. 2003. Grammatical metaphor, controlled language and machine translation. *IN: European association for machine translation: international workshop*, May 2003. Dublin: Dublin City University, pp95-104.
- Nyberg, E.H. and Mitamura, T. 1996. Controlled language and knowledge-based machine translation: principles and practice. *IN: Proceedings of the first controlled language application workshop*, March 1996. Leuven: Centre for Computational Linguistics, pp74-83.
- O'Brien, S. 2003. Controlling controlled English: an analysis of several controlled language rule sets. *IN: European association for machine translation: international workshop*, May 2003. Dublin: Dublin City University, pp105-114.
- O'Brien, S. and Roturier, J. 2007. How portable are controlled languages rules: a comparison of two empirical MT studies [Online]. Available from: <http://www.mt-archive.info/MTS-2007-OBrien.pdf> [Accessed 2 September 2008].
- Oppenheim, A.N. 1983. *Questionnaire design and attitude measurement*. 10<sup>th</sup> ed. London: Heinemann.
- Puurtinen, T. 1995. *Linguistic acceptability in translated children's literature*. Joensuu: University of Joensuu.
- Redish, J. 2000. Usability testing reveals more than readability formulas reveal. *ACM journal of computer documentation*. 24 (3), pp132-137.
- Reuther, U. 1998. Controlling language in an industrial application. *IN: Proceedings of the second international workshop on controlled language applications*. May 1998. Pennsylvania: Carnegie Mellon University, pp174-184.
- Reuther, U. 2003. Two in one – can it work? Readability and translatability by means of controlled language. *IN: European association for machine translation: international workshop*, May 2003. Dublin: Dublin City University, pp124-132.

Roturier, J. 2006. *An investigation into the impact of controlled English rules on the comprehensibility, usefulness and acceptability of machine-translated technical documentation for French and German users*. PhD Thesis. Dublin City University.

Schriver, K. 2000. Readability formulas in the new millennium: what's the use? *ACM journal of computer documentation*. 24 (3), pp138-140.

Shubert, S.K., Spyridakis, J.H., Holmback, H.K. and Coney, M.B. 1995. The comprehensibility of Simplified English in procedures. *Journal of technical writing and communication*. 25 (4), pp347-369.

Silverman, D. 2006. *Interpreting qualitative data: methods for analyzing talk, text, and interaction*. 3rd ed. London; Thousand Oaks, California: SAGE Publications.

Spaggiari, L., Beaujard, F. and Cannesson, E. 2003. A controlled language at airbus. *IN: European association for machine translation: international workshop*, May 2003. Dublin: Dublin City University, pp151-159.

Wheeler, M., Ewers, M. and Buonanno, J. 2003. Different rates of forgetting following study versus test trials. *Memory*. 11 (6), pp571-580.

Wojcik, R., Holmback, H. and Hoard, J. 1998. Boeing technical English: an extension of AECMA SE beyond the aircraft maintenance domain. *IN: Proceedings of the second international workshop on controlled language applications*. May 1998. Pennsylvania: Carnegie Mellon University, pp114-123.

(Word count 12,563 words)

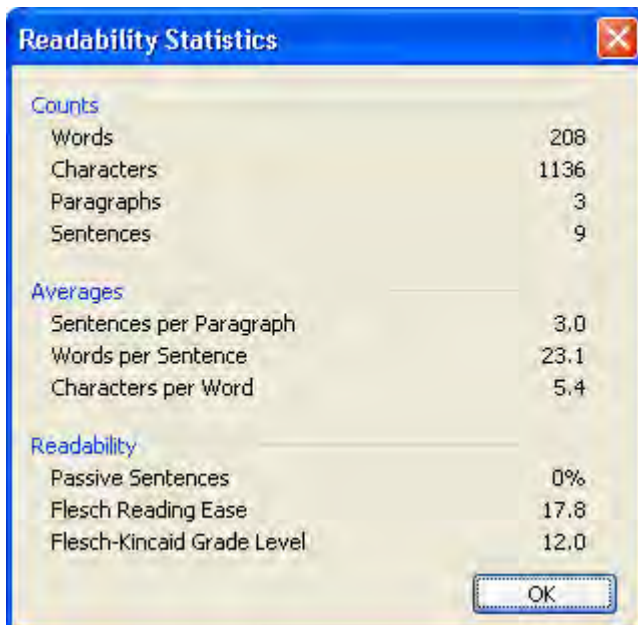
## Appendix A: NCL passages and their readability scores

### 1. 'Norton' passage: predicted to be the most difficult of the three to read

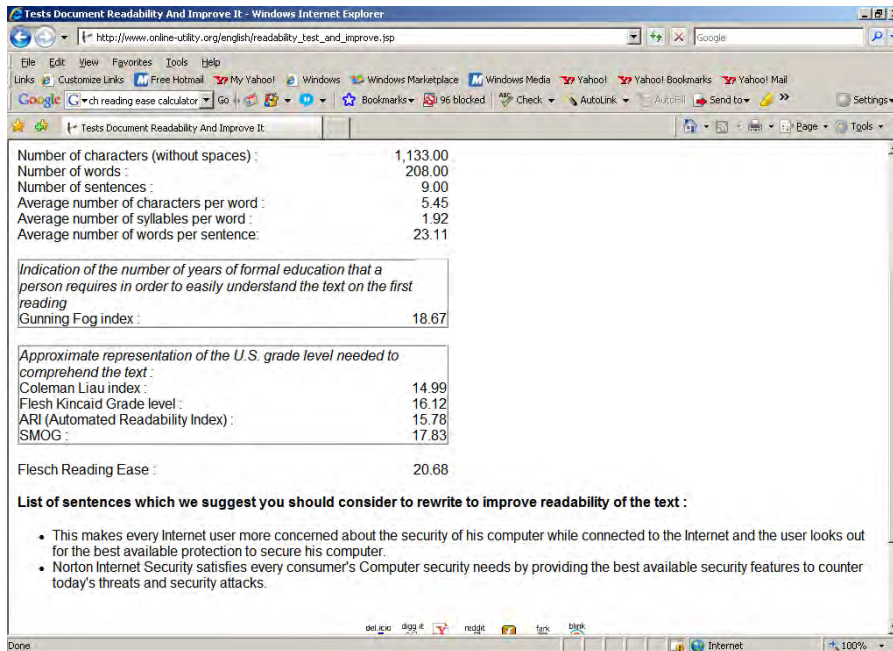
Norton Internet Security 2007 is the tenth version in the product line. While today's Internet provides a wealth of information and resources, it is also a gateway for threats and hackers to enter or exploit a user's computer. This makes every Internet user more concerned about his PC security and the user looks out for the best available security product in the market. Norton Internet Security 2007 satisfies every consumer's PC security needs by providing the best available security features to counter today's threats and security attacks.

With the growing dependency on Internet which provides instant access to a range of information and resources that could be of great aid. A need for protection against malicious content which enters and exploits a computers security while on the internet has increased.

This makes every Internet user more concerned about the security of his computer while connected to the Internet and the user looks out for the best available protection to secure his computer. Norton Internet Security satisfies every consumer's Computer security needs by providing the best available security features to counter today's threats and security attacks. Norton internet Security 2007 the tenth version of this product continues to provide optimum security as its predecessors, through its new enhanced features.



Readability Statistics	
<b>Counts</b>	
Words	208
Characters	1136
Paragraphs	3
Sentences	9
<b>Averages</b>	
Sentences per Paragraph	3.0
Words per Sentence	23.1
Characters per Word	5.4
<b>Readability</b>	
Passive Sentences	0%
Flesch Reading Ease	17.8
Flesch-Kincaid Grade Level	12.0



## 2. 'HijackThis' passage: predicted to be the next most difficult of the three to read

### HijackThis

This section will provide you a brief overview on the HijackThis tool. After you complete this section, you will be able to do the following:

- Describe the usage of the HijackThis tool
- Analyze the results / logs generated by the tool
- Troubleshoot a relevant issue using the tool

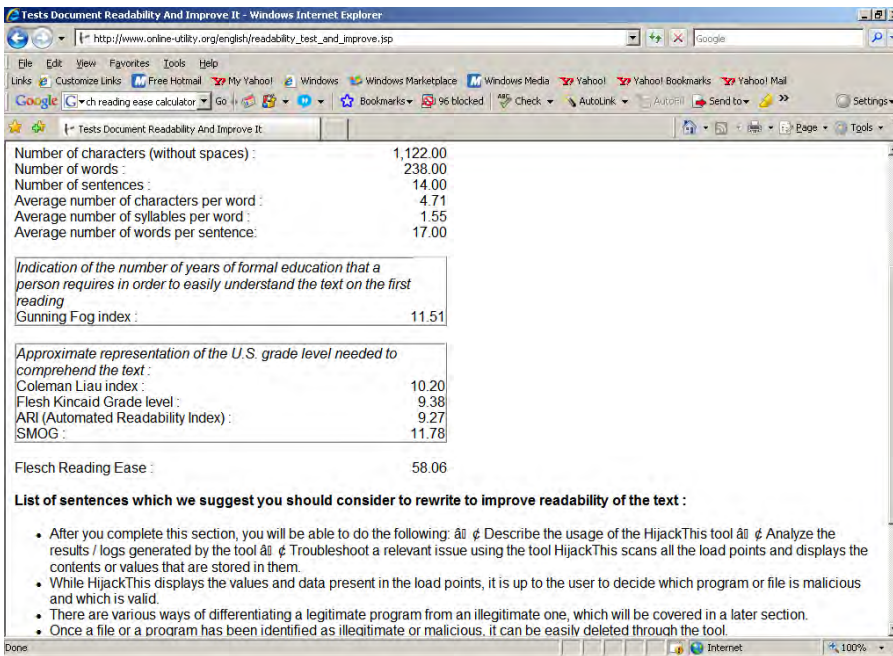
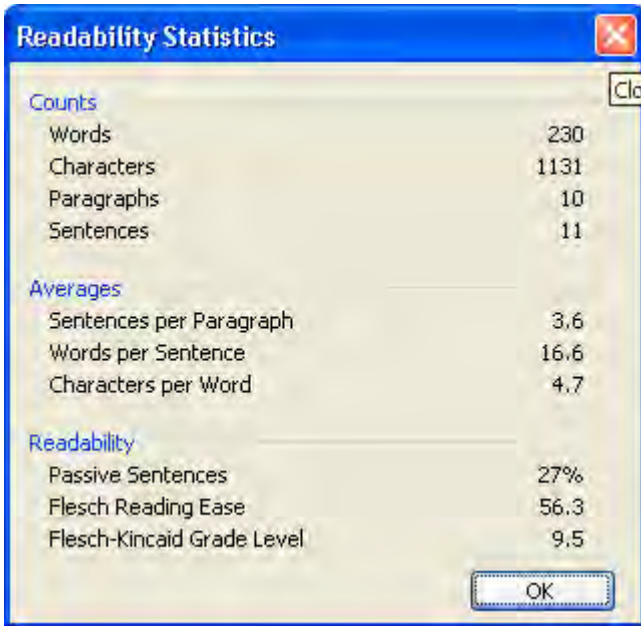
HijackThis scans all the load points and displays the contents or values that are stored in them. It also shows the Processes that run in the background when the tool is run. While HijackThis displays the values and data present in the load points, it is up to the user to decide which program or file is malicious and which is valid. Once a file or a program has been identified as illegitimate or malicious, it can be easily deleted through the tool. The HijackThis tool itself cannot differentiate between a legitimate and an illegitimate program. There are various ways of differentiating a legitimate program from an illegitimate one, which will be covered in a later section. Let's now have a look at the tool itself.

#### Obtaining the tool

HijackThis can be downloaded from the following link:

<http://www.hijackthis.de> Note: Extract the downloaded zip file and save HijackThis.exe in a folder.

Double-clicking on the tool should open a screen with several options. To analyze all load points and running tasks, click on the "Do a System Scan and Save log file" button.



### 3. 'Shared' passage: predicted to be the least difficult of the three to read

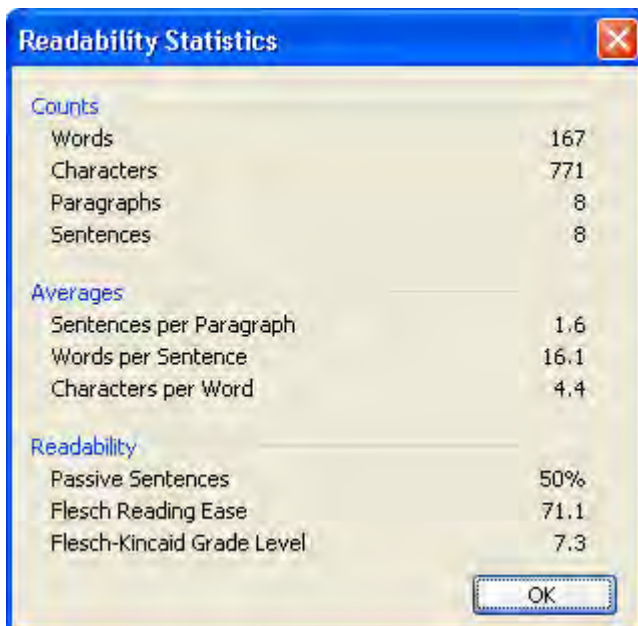
## Shared Folders

This scanner checks if there are shared resources on a user's system and provides an option to "fix" or "unshare" them. It ideally looks for the following criteria:

- Global shares must not be enabled.
- System folders should not be shared.

After a scan is run, and the user is presented with a list of shares that will be closed, here's the behavior of what is presented to the user:

- If there are any Global Shares (i.e. C\$, D\$, etc.) or System-Folder Shares (i.e. C:\Windows or C:\Windows\System, etc.), then no user-created shares will be listed/closed. This is so that the most-critical shares are closed first.
- If there are less than 5 user-created shares on the system, then nothing will be listed/closed.
- If there are 5 or more user-created shares on the system, they will all be listed in the details dialog, but the default action will be set to "No Action". This is to avoid any accidental decision by the user.



The screenshot shows a dialog box titled "Readability Statistics" with a close button in the top right corner. The dialog is divided into three sections: "Counts", "Averages", and "Readability". Each section contains a list of metrics and their corresponding values.

Counts	
Words	167
Characters	771
Paragraphs	8
Sentences	8

Averages	
Sentences per Paragraph	1.6
Words per Sentence	16.1
Characters per Word	4.4

Readability	
Passive Sentences	50%
Flesch Reading Ease	71.1
Flesch-Kincaid Grade Level	7.3

An "OK" button is located at the bottom right of the dialog box.

Tests Document Readability And Improve It - Windows Internet Explorer

http://www.online-utility.org/english/readability\_test\_and\_improve.jsp

File Edit View Favorites Tools Help

Links Customize Links Free Hotmail My Yahoo! Windows Windows Marketplace Windows Media Yahoo! Yahoo! Bookmarks Yahoo! Mail

Google ch reading ease calculator Go 96 blocked Check AutoLink Send to Settings

Tests Document Readability And Improve It

Number of characters (without spaces) :	758.00
Number of words :	176.00
Number of sentences :	14.00
Average number of characters per word :	4.31
Average number of syllables per word :	1.49
Average number of words per sentence :	12.57

*Indication of the number of years of formal education that a person requires in order to easily understand the text on the first reading*

Gunning Fog index :	7.30
---------------------	------

*Approximate representation of the U.S. grade level needed to comprehend the text :*

Coleman Liau index :	7.18
Flesh Kincaid Grade level :	6.88
ARI (Automated Readability Index) :	5.14
SMOG :	8.86

Flesch Reading Ease : 68.14

**List of sentences which we suggest you should consider to rewrite to improve readability of the text :**

- After a scan is run, and the user is presented with a list of shares that will be closed, here's the behavior of what is presented to the user:
  - If there are any Global Shares (i
  - If there are 5 or more user-created shares on the system, they will all be listed in the details dialog, but the default action will be set to "No Action".
  - Shared Folders This scanner checks if there are shared resources on a user's system and provides an option to fix or unshare them.
  - If there are less than 5 user-created shares on the system, then nothing will be listed/closed.

Done Internet 100%

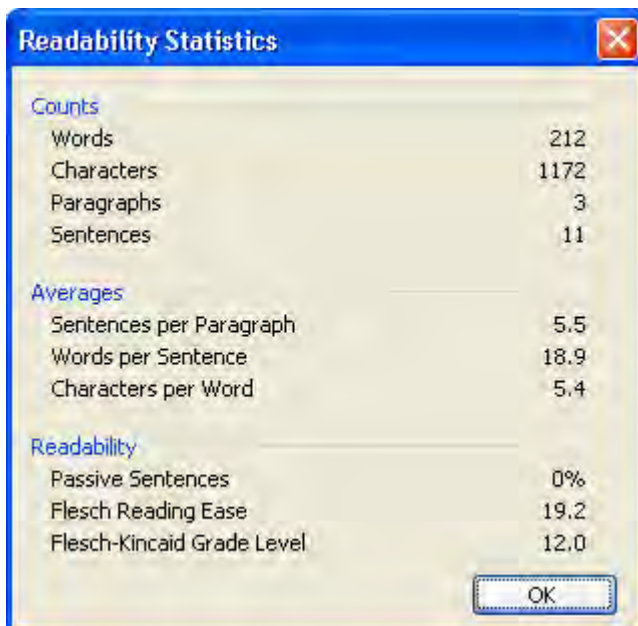
## Appendix B: CL passages and their readability scores

### 1. 'Norton' passage: predicted to be the most difficult of the three to read

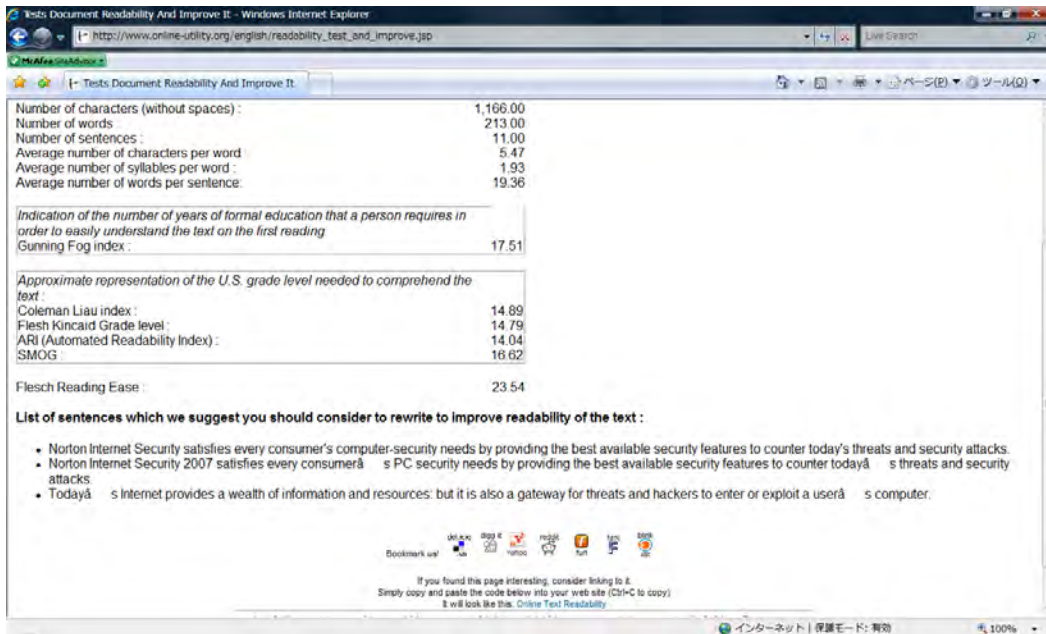
#### **Norton Internet Security 2007**

Norton Internet Security 2007 is the tenth version in the product line. Today's Internet provides a wealth of information and resources: but it is also a gateway for threats and hackers to enter or exploit a user's computer. This fact makes every Internet user more concerned about PC security, and the user looks out for the best available security product in the market. Norton Internet Security 2007 satisfies every consumer's PC security needs by providing the best available security features to counter today's threats and security attacks.

These security features can be of great aid as more people grow dependent on the Internet to access a range of information and resources instantly. A need for protection against the malicious content which can enter and exploit a computer's security while online has increased. This need makes every Internet user more concerned about their computer security while online. The user looks out for the best available protection to achieve this computer security. Norton Internet Security satisfies every consumer's computer-security needs by providing the best available security features to counter today's threats and security attacks. Norton internet Security 2007 is the tenth version of this product. As with its predecessors, it continues to provide optimum security through its new, enhanced features.



Readability Statistics	
<b>Counts</b>	
Words	212
Characters	1172
Paragraphs	3
Sentences	11
<b>Averages</b>	
Sentences per Paragraph	5.5
Words per Sentence	18.9
Characters per Word	5.4
<b>Readability</b>	
Passive Sentences	0%
Flesch Reading Ease	19.2
Flesch-Kincaid Grade Level	12.0
OK	



## 2. 'HijackThis' passage: predicted to be the next most difficult of the three to read

### HijackThis

This section provides you with a brief overview on the HijackThis tool. After you complete this section, you should be able to do the following:

- Describe the usage of the HijackThis tool
- Analyze the results or logs that the tool generates
- Troubleshoot a relevant issue using the tool

HijackThis scans all the load points and displays the contents or values that are stored in them. It also shows the Processes that run in the background when the tool is run. HijackThis displays the values and data present in the load points. But it is up to the user to decide which program or file is malicious and which program or file is valid. Once a file or a program has been identified as illegitimate or malicious, it can be easily deleted through the tool. The HijackThis tool itself cannot differentiate between a legitimate and an illegitimate program. A legitimate program can be differentiated from an illegitimate program in various ways: these ways are covered in a later section. Let's now have a look at the tool itself.

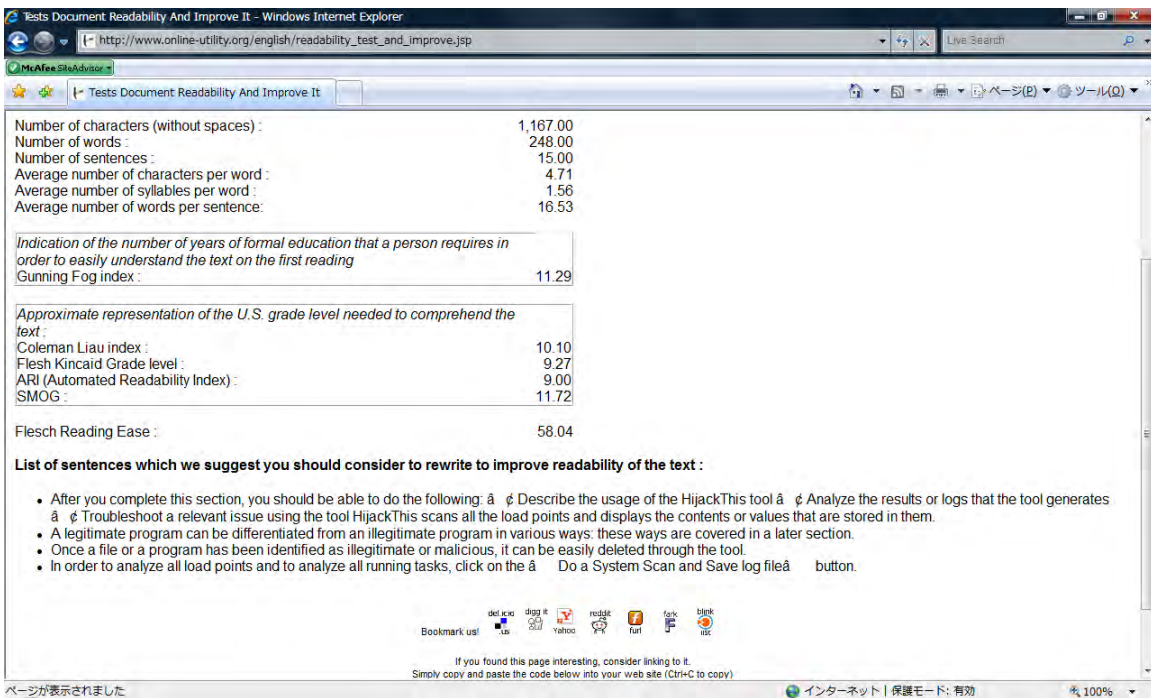
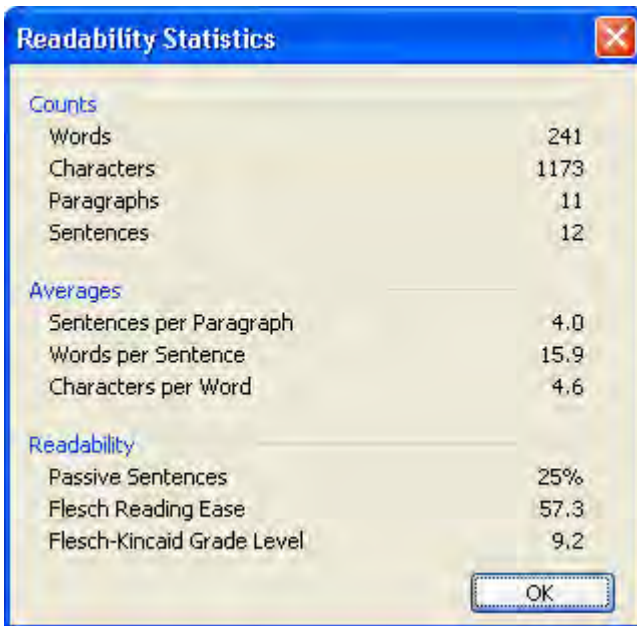
#### Obtaining the tool

HijackThis can be downloaded from the following link:

<http://www.hijackthis.de>

Note: Extract the downloaded zip file and save "HijackThis.exe" in a folder.

If you double-click on the tool, a screen with several options should open. In order to analyze all load points and to analyze all running tasks, click on the "Do a System Scan and Save log file" button.



### 3. 'Shared' passage: predicted to be the least difficult of the three to read

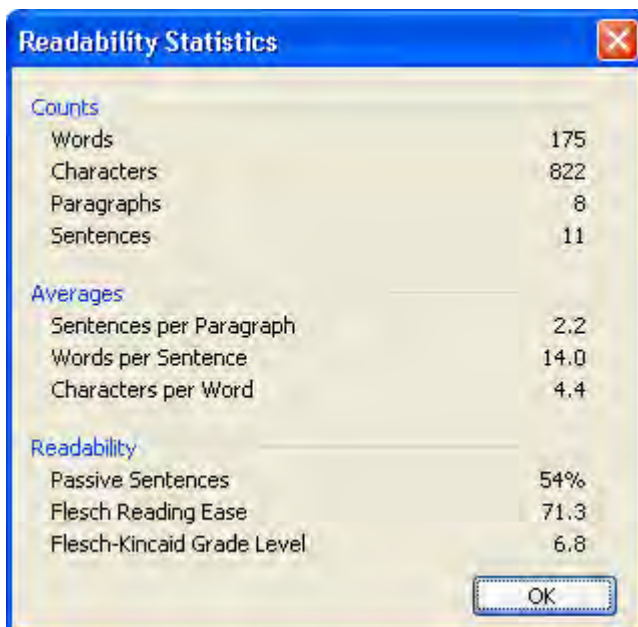
## Shared Folders

This scanner checks if there are shared resources on a user's system and provides an option to "fix" or "unshare" them. It ideally looks for the following criteria:

- Global shares must not be enabled.
- System folders should not be shared.

After a scan is run, the system presents the user with a list of shares to be closed. Here's the behavior of what is then presented to the user:

- If there are Global Shares (i.e. C\$, D\$, etc.), no user-created shares are listed or closed. Also, if there are System-Folder Shares (i.e. C:\Windows, C:\Windows\System, etc.), no user-created shares are listed or closed. These steps mean that the most-critical shares get closed first.
- If there are less than 5 user-created shares on the system, nothing is listed or closed.
- If there are 5 or more user-created shares on the system, they are all listed in the details dialog. But the default action is set to "No Action" in this case. This setting is to avoid any accidental decision by the user.



The screenshot shows a dialog box titled "Readability Statistics" with a close button (X) in the top right corner. The dialog is divided into three sections: "Counts", "Averages", and "Readability". Each section contains a list of metrics and their corresponding values.

Counts	
Words	175
Characters	822
Paragraphs	8
Sentences	11

Averages	
Sentences per Paragraph	2.2
Words per Sentence	14.0
Characters per Word	4.4

Readability	
Passive Sentences	54%
Flesch Reading Ease	71.3
Flesch-Kincaid Grade Level	6.8

An "OK" button is located at the bottom right of the dialog box.

Tests Document Readability And Improve It - Windows Internet Explorer  
 http://www.online-utility.org/english/readability\_test\_and\_improve.jsp

Number of characters (without spaces) : 805.00  
 Number of words : 187.00  
 Number of sentences : 17.00  
 Average number of characters per word : 4.30  
 Average number of syllables per word : 1.48  
 Average number of words per sentence : 11.00








*Indication of the number of years of formal education that a person requires in order to easily understand the text on the first reading*  
 Gunning Fog index : 6.11

*Approximate representation of the U.S. grade level needed to comprehend the text :*  
 Coleman Liau index : 6.83  
 Flesh Kincaid Grade level : 6.18  
 ARI (Automated Readability Index) : 4.35  
 SMOG : 7.79

Flesch Reading Ease : 70.35

**List of sentences which we suggest you should consider to rewrite to improve readability of the text :**

- Shared Folders This scanner checks if there are shared resources on a user's system and provides an option to fix or unshare them.
- If there are 5 or more user-created shares on the system, they are all listed in the details dialog.
- Here's the behavior of what is then presented to the user: If there are Global Shares (i.e., shares that are not user-created), they are listed first.
- If there are less than 5 user-created shares on the system, nothing is listed or closed.
- It ideally looks for the following criteria: Global shares must not be enabled.

Bookmark us!       

If you found this page interesting, consider linking to it.  
 Simply copy and paste the code below into your web site (Ctrl+C to copy)

ページが表示されました インターネット | 保護モード: 有効 100%

## Appendix C: Sample questionnaire

### Sample questionnaire (Stages 1 and 2 combined)

Stage

Group

No.

Name: \_\_\_\_\_

(This will not be used in the final report)

## **1. Introduction**

---

In this questionnaire, I want to ask you about readability.

Take the famous phrase "*To be or not to be*". It is extremely easy to read; even a young child can do it. So, we say that such a sentence has very high readability.

However, can we say that this phrase is easy to understand or comprehend? A school student, an English teacher and an actor may all understand this phrase in very different ways.

Also, can we say it is easy to translate into another language? And would all translators do it in the same way?

In short, to say something is *easy to read* (or has *high readability*) is not necessarily to say that it is easy to understand or to translate.

When you are filling out this questionnaire, please remember that I want to know what you think about **readability**. The questions are not about how well you understand the underlying subject matter; they're about how easy or difficult the passages are to read.

Also, in this questionnaire, I sometimes talk about *training documents*. I intend *training documents* to mean written materials that help you to perform tasks better and that help you to know more about your working role.

With this these points in mind, please move on to the first question.

## **2. Training documents and readability**

In this section, I ask you about your thoughts on training documents and readability in general.

### **Question 1a**

Please read the following statement and put a tick under the option which best describes your attitude.

	Strongly agree	Agree	Disagree	Strongly disagree	No Opinion
It is important for training documents to be easy to read.					

### **Question 1b**

If you agreed or disagreed with this statement, why do you feel this way?

---

---

---

---

**Question 2a**

Please tick the option which best describes the way you read training documents. If you want to choose more than one option, please give more detail in 2b below.

**Skim** (*reading the headings and first sentences of paragraphs*)

**Scan** (*search through the material for a specific purpose or word*)

**In-depth** (*try to understand the concepts and arguments that the text contains*)

**Question 2b**

---

---

---

---

**Question 3**

Please match each concept in Column A to the most appropriate definition in Column B. Draw a line from the concept to the definition to indicate your choice.

**Col. A: Concept**

Comprehensibility

Readability

Acceptability

**Col. B: Definition**

The degree to which a translation conforms to the reader's expectations and preferences.

The ability to produce a message in a way that the recipient can discern the appropriate meaning.

The characteristic of a text that indicates the clarity of its display and its message.

### 3. Reading passages and evaluating readability

Now I'm going to ask you to read three passages. Please note down exactly in the spaces provided the time that you start and stop reading each passage: there is no rush. Please read at a comfortable speed. After each passage, I'll ask you a question; **remember** that this question will be about **readability**.

## Start Time

H	H	:	M	M	:	S	S

Passages appeared here, but have not been included in this sample questionnaire.

## Finish Time

H	H	:	M	M	:	S	S

#### Question 4

Please select the option which best describes the passage you have just read:

*It was **fairly easy** to read*

*It was **somewhat difficult** to read*

*It was **very challenging** to read*

## **4. Reading passages and highlighting difficulty**

Now I'm going to show you exactly the same three passages: they are printed out again below. This time, as you read through the passages, underline the parts of the text that you find **difficult to read**, if any.

Passages appeared here, but have not been included in this sample questionnaire.

## **5. Retaining vocabulary and concepts**

In this last section I want to explore the content that we retain after reading a passage. Please DO NOT look back to the previous pages to try to answer this question.

### **Question 6a**

You have now read three passages about Norton Internet Security 2007. Have you ever read these passages before today?

Yes

No

I don't remember

**Question 6b**

Did any of the following words appear in the three passages? If so, please tick as appropriate.

Safety

Shares

Tool

Device

Security

Allocation

Extra Question only in Stage 2 Questionnaire

## **5. Comparing versions of the passages**

This last section compares two versions of each passage: one version is from the last session, and one is from today. I want to know which version was easier to read.

Passages appeared here, but have not been included in this sample questionnaire.

**Question 5a**

Please tick the most appropriate choice:

I think the **version from the last time** is easier to read.

I think the **version from today** is easier to read.

I **don't know**.

## **6. Participant profile**

The questions on this page are optional. Answers given here will just be used to build a general profile of all participants: no individual participant details will appear in the final report.

### **Question 6.1**

What is your occupation?

---

### **Question 6.2**

Do you use a computer in this occupation?

Yes

No

### **Question 6.3**

What is your age?

---

### **Question 6.4**

Please tick the levels of education that you have completed:

Post-primary level (e.g. Leaving Certificate or its equivalent)

Undergraduate level (e.g. Bachelors Degree)

Postgraduate level (e.g. Masters Degree, PhD)

**Question 6.5a**

Are you a native speaker of English?

Yes

No

**Question 6.5b**

If no, please put an X through the number which best describes your reading ability in English.

1= Beginner Level

10= Native-speaker Level

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

# Appendix D: Tables of raw data from questionnaires

**Stage 1: raw data compiled from questionnaires**  
**Participants in column1 (A = Symantec, B = Control)**  
**Questions numbered as per questionnaire**

Question	H - High skills	M - Noncon	S - Shared	MCL words	223	202	164	6.2	6.3	6.4	6.5	6.5b	6.5c	6.5d	6.5e	6.5f	6.5g	6.5h	6.5i	6.5j	6.5k	6.5l	6.5m	6.5n	6.5o	6.5p	6.5q	6.5r	6.5s	6.5t	6.5u	6.5v	6.5w	6.5x	6.5y	6.5z					
A1 (Participant)	MCL	Strongly agree	understand, if trainer scan, in-dept depend what	Correct	181	0.190	Correct: som	58	0.287	Incorrect: fail	43	0.262	Correct: fail	No	2	1	Technical Product Support Engineer	Yes	45	1	No	7	M																		
A2	MCL	Strongly agree	understand concept in-depth	als	Correct	120	0.524	Incorrect: fail	120	0.534	Incorrect: som	60	0.366	Correct: fail D.R.	1	0	Trainer/ Project Coordinator	Yes	33	3	No	10	F																		
A3	CL	Strongly agree	familiarizing with prc scan, in-dept scan for quick	Correct	100	0.417	Correct: som	70	0.347	Incorrect: fail	50	0.287	Correct: fail	No	3	2	Localization coordinator	Yes	30	3	No	8	F																		
A4	CL	Agree	understand, absorb, scan, in-dept depend press	Correct	80	0.333	Correct: som	70	0.347	Incorrect: fail	110	0.632	Incorrect: som	No	3	1	Language-Coordinator	Yes	33	3	Yes	ns	F																		
A5	CL	Strongly agree	better understanding in-depth	als	Incorrect (2/3)	60	0.250	Incorrect: fail	120	0.534	Incorrect: fail	60	0.345	Correct: fail	No	2	0	Tranahion Coordinator	Yes	26	2	Yes	ns	F																	
A6	MCL	Agree	Supposed to help th in-depth	als	Correct	105	0.453	Correct: som	70	0.347	Incorrect: fail	60	0.366	Incorrect: som	No	3	0	Project Coordinator	Yes	34	3	No	9	F																	
A7	CL	Strongly agree	need feel understand in-depth	als	Correct	150	0.625	Correct: som	120	0.534	Incorrect: fail	120	0.630	Incorrect: som	No	3	0	Technical Transition Support	Yes	30	2	Yes	ns	F																	
A8	MCL	Strongly agree	if flows correctly as dlim, in-dept dlim to find if	Correct	63	0.275	Correct: som	58	0.287	Incorrect: som	46	0.230	Correct: fail	No	3	1	Principal product support engineer (s	Yes	29	3	No	10	M																		
A9	CL	Strongly agree	to focus on inform scan, in-dept scan for impo	Correct	72	0.300	Correct: som	42	0.208	Incorrect: fail	43	0.247	Incorrect: som	No	3	1	Software support Engineer	Yes	33	2	Yes	ns	M																		
A10	MCL	Strongly agree	if not easy danger s/ in-depth	als	Incorrect (2/3)	35	0.163	Incorrect: fail	33	0.460	Incorrect: som	83	0.506	Incorrect: som	No	2	1	Technical Engineer	Yes	49	ns	No	6	M																	
A11	MCL	Strongly agree	focus on content scan, in-dept scan to find if	Incorrect (2/3)	143	0.624	Incorrect: fail	20	0.093	Incorrect: fail	70	0.427	Incorrect: som	No	1	1	Project coordinator	Yes	42	2	No	10	F																		
A12	CL	Strongly agree	improve knowledge in-depth	als	Incorrect (2/3)	50	0.208	Incorrect: fail	100	0.435	Correct: som	30	0.182	Correct: fail	No	0	0	Technical support	Yes	43	1	No	9	M																	
B1	CL	Agree	better choice under dlim, in-dept find relevant	Correct	82	0.383	Correct: som	71	0.351	Incorrect: fail	57	0.269	Incorrect: very dal	3	2	Teacher	Yes	38	3	Yes	ns	F																			
B2	CL	Strongly agree	if not, delete pumps scan	Correct	120	0.500	Correct: som	60	0.237	Incorrect: fail	76	0.437	Incorrect: very dal	3	1	Finance	Yes	32	2	Yes	ns	F																			
B3	CL	Strongly agree	clear steps easy to f dlim, scan	dlim first, sco	Correct	82	0.383	Incorrect: fail	112	0.514	Incorrect: som	69	0.397	Correct: fairly easy	1	1	Graduate student	Yes	28	3	Yes	ns	M																		
B4	MCL	Strongly agree	not waste time scan	Correct	70	0.306	Correct: som	64	0.317	Incorrect: fail	46	0.293	Incorrect: somewhat	3	1	Graduate student	Yes	32	2	Yes	ns	M																			
B5	CL	Strongly agree	hard understand dlim, in-dept if not interest	Correct	134	0.593	Incorrect: fail	53	0.262	Incorrect: fail	152	0.814	Incorrect: somewhat	2	1	Commis dtd	No	29	2	Yes	ns	M																			
B6	MCL	Strongly agree	understand 100% not dlim, in-dept it depends on	Correct	160	0.689	Correct: som	115	0.563	Incorrect: fail	75	0.457	Correct: fairly easy	3	1	Graphic designer	Yes	32	3	No	8	F																			
B7	MCL	Strongly agree	easy understand scan	Correct	55	0.240	Correct: som	45	0.223	Incorrect: fail	35	0.213	Correct: fairly easy	1	1	Graduate student	Yes	23	3	No	7	F																			
B8	MCL	Strongly agree	less interest mis: fail scan, in-dept identify impo	Correct	89	0.432	Correct: som	100	0.435	Incorrect: fail	68	0.415	Incorrect: somewhat	3	1	Saler representative	Yes	37	3	No	9	F																			
B9	MCL	Strongly agree	should make things c in-depth	Correct	60	0.282	Incorrect: fail	75	0.371	Incorrect: fail	50	0.305	Correct: fairly easy	3	0	Graduate student	Yes	ns	3	No	ns	F																			
B10	CL	Agree	help perform tasks t dlim, scan, in depends on k	Correct	91	0.379	Correct: som	50	0.248	Incorrect: fail	62	0.395	Correct: fairly easy	2	1	Student dall student	No	21	2	No	8	M																			
B11	MCL	Strongly agree	propose to get som scan, in-dept scan to secret	Correct	80	0.349	Incorrect: fail	30	0.143	Incorrect: fail	115	0.701	Incorrect: somewhat	3	0	Teacher	Yes	30	3	No	8	F																			
B12	CL	Agree	enjoy not scan if rec scan, in-dept scan first in-d	Incorrect (2/3)	65	0.271	Incorrect: fail	25	0.124	Incorrect: som	40	0.230	Correct: fairly easy	3	1	Graduate student	Yes	22	3	No	10	F																			

**Stage 2: raw data from questionnaires**

**Participants in column 1 (A = Symantec, B = Control)  
Questions numbered as per questionnaire**

Stage 2		NCL words		229		202		164							
Question	Time H	Time p/w H	2a	Time N	Time p/w N	2b	Time S	Time p/w S	2c	4: retained	4: noise	5: Preference	Morton	Hipack	Shared
A1(Participant)	48	0.200	Correct: somewhat difficult	53	0.282	Incorrect: fairly easy	42	0.241	Correct: fairly easy	3	1	CL			
A2	65	0.271	Incorrect: fairly easy	105	0.520	Incorrect: fairly easy	80	0.450	Correct: fairly easy	3	0	CL			
A3	40	0.175	Incorrect: fairly easy	60	0.297	Incorrect: somewhat difficult	25	0.192	Correct: fairly easy	3	1	CL			
A4	58	0.253	Incorrect: fairly easy	65	0.322	Incorrect: somewhat difficult	50	0.305	Correct: fairly easy	3	1	CL			
A5	30	0.181	Incorrect: fairly easy	125	0.619	Incorrect: fairly easy	50	0.305	Correct: fairly easy	3	1	CL			
A6	45	0.198	Incorrect: fairly easy	60	0.297	Incorrect: fairly easy	30	0.172	Correct: fairly easy	3	1	CL			
A7	50	0.218	Incorrect: fairly easy	120	0.594	Incorrect: somewhat difficult	90	0.549	Incorrect: somewhat difficult	3	1	DK			
A8	70	0.292	Correct: somewhat difficult	65	0.322	Incorrect: somewhat difficult	45	0.293	Correct: fairly easy	3	1	CL			
A9	55	0.240	Correct: somewhat difficult	40	0.198	Incorrect: fairly easy	27	0.165	Correct: fairly easy	3	0	NCL			
A10	80	0.333	Incorrect: fairly easy	75	0.371	Incorrect: somewhat difficult	60	0.345	Correct: fairly easy	3	1	NCL			
A11	90	0.375	Incorrect: fairly easy	65	0.322	Incorrect: fairly easy	65	0.374	Incorrect: somewhat difficult	3	0	NCL			
A12	35	0.193	Correct: somewhat difficult	40	0.198	Correct: very challenging	30	0.193	Correct: fairly easy	0	0	NCL			
B1	72	0.314	Correct: somewhat difficult	74	0.366	Incorrect: fairly easy	53	0.323	Incorrect: somewhat difficult	3	2	CL			
B2	75	0.328	Incorrect: fairly easy	70	0.347	Incorrect: somewhat difficult	50	0.305	Incorrect: very challenging	2	1	CL			
B3	50	0.218	Incorrect: fairly easy	82	0.406	Incorrect: somewhat difficult	37	0.226	Correct: fairly easy	1	1	NCL			
B4	83	0.346	Correct: somewhat difficult	84	0.416	Incorrect: fairly easy	53	0.305	Incorrect: somewhat difficult	2	1	CL			
B5	41	0.179	Incorrect: fairly easy	57	0.282	Incorrect: fairly easy	48	0.293	Correct: fairly easy	3	1	DK			
B6	122	0.508	Incorrect: fairly easy	65	0.322	Incorrect: fairly easy	91	0.523	Incorrect: somewhat difficult	3	1	DK			
B7	48	0.200	Incorrect: fairly easy	40	0.198	Incorrect: fairly easy	30	0.172	Correct: fairly easy	2	0	CL			
B8	80	0.333	Incorrect: fairly easy	22	0.109	Incorrect: fairly easy	60	0.345	Correct: fairly easy	3	2	NCL			
B9	50	0.208	Incorrect: fairly easy	50	0.248	Incorrect: fairly easy	35	0.201	Correct: fairly easy	3	0	CL			
B10	58	0.253	Incorrect: fairly easy	57	0.282	Incorrect: somewhat difficult	55	0.335	Correct: fairly easy	2	2	NCL			
B11	90	0.375	Incorrect: fairly easy	90	0.446	Incorrect: fairly easy	120	0.630	Incorrect: somewhat difficult	3	0	CL			
B12	60	0.262	Incorrect: fairly easy	40	0.198	Incorrect: fairly easy	30	0.183	Correct: fairly easy	1	1	NCL			