College of Science and Health Theses and Dissertations

College of Science and Health

Fall 11-22-2022

How the Readability Level of Prior Text Impacts Comprehension of Subsequent Text

Benjamin Jerry Dumont
DePaul University, bdumont@depaul.edu

Follow this and additional works at: https://via.library.depaul.edu/csh_etd

Part of the Psychology Commons

How the Readability Level of Prior Text Impacts Comprehension of Subsequent Text

A Thesis

Presented in

Partial Fulfillment of the

Requirements for the Degree of

Master of Science

By

Ben Dumont

May 2022

Department of Psychology

College of Science and Health

DePaul University

Chicago, Illinois

Thesis Committee

Jessica Choplin, PhD, Chair

David Allbritton, PhD

Acknowledgments

I would like to thank my faculty advisor, Dr. Jessica Choplin, for her mentorship and support throughout the thesis process. I would also like to thank my thesis reader, Dr. David Allbritton, for his insight with planning and analyzing the final study, on top of his overall advice. Without their assistance and guidance, this work would not have been possible. I would also like to thank my partner, Sam, for standing by me in my academic career and always being there to support and push me. Jason Snyder, my statistics TA, also deserves a thank you for taking the time to answer my statistics questions. Lastly, I would like to thank my parents, Phil and Lisa, for always encouraging my pursuit of knowledge. Without their influence to work hard in school and pursue what I enjoy, I would not have reached this point in my academic career.

Biography

The author was born in St Paul, Minnesota, on June 30th, 1997. He graduated from River Falls

High School, in River Falls, Wisconsin in 2016. He received his Bachelor of Science in

Psychology and Communication Studies from University of Wisconsin – La Crosse in 2020.

Table of Contents

List of Tables

Abstract

The purpose of this thesis research is to explore how the difficulty of text impacts the ability to process subsequent text. The basic idea is that difficult text might deplete cognitive or attentional resources, creating difficulties for readers as they try to continue reading afterwards. A pilot study was conducted to gather preliminary data prior to the two final studies. High (easy to read) and low (difficult to read) readability passages were created and presented prior to a target passage. The results from the pilot study suggested that there may be differences in comprehension based on the readability and difficulty level of preceding text. Study one investigated how the length of the preceding passage impacted comprehension of subsequent text. A 2 readability (easy vs hard) x 2 length (short vs long) ANOVA failed to find any significant main effects or a significant interaction. Study two investigated whether differences in working memory capacity might make some people more susceptible to the effects of difficulty on comprehension on subsequent text. An ANCOVA failed to find a significant effect of readability on comprehension of the subsequent text after controlling for verbal working memory scores. Working memory scores were also not a significant predictor of comprehension scores.

How the Readability Level of Prior Text Impacts Comprehension of Subsequent

Text

Taxi driver Mohammed Hoque signed a contract to buy a taxi medallion for $50,000 dollars (Rosenthal, 2019). What he did not notice or read was that the contract he signed required him to pay $1.7 million, which he could not afford on the money he made from his taxi fares. He ended up going bankrupt since he drained his accounts to pay for the medallion. This was happening across New York in the taxi industry at the time, which found "drivers trapped in exploitative loans … that required them to forfeit their legal rights" (Rosenthal, 2019, p.4). There are several potential reasons why Mohammed may not have understood what he was signing. He may have just skimmed the contract, instead of reading it carefully. Maybe he tried to read it but was unable to fully understand the complicated legal language that was used, or he may have used too many cognitive resources early on trying to understand the opening of the contract, which may have made reading and comprehending the subsequent pages much more difficult. This can be a problem and can contribute to a lack of comprehension and, in drastic circumstances, financial insecurity and bankruptcy. The purpose of the studies reported was to understand what caused Mohammed and hundreds of thousands other consumers like him to misunderstand the terms of the contracts they sign. These studies investigated whether text readability, text position, length, and cognitive resource consumption can have an impact on reader comprehension of what follows, especially in the content of legal documents. Might difficult text in his contract have caused him difficulty such that he was less able to process subsequent legal clauses?

Previous research has identified several factors that can affect comprehension, including script activation, attention, optimum effort, and working memory (Abelson, 1981; Smith & Houston; Weaver & Bryant, 1995; Elsayyad, Everatt, Mortimore, & Haynes, 2017). These factors may have played a role in why Mohammed failed to understand that the contract he signed required him to pay much more, $1.7 million, than he could afford for the taxicab medallion. Each of the factors will be discussed and explained how they might have caused difficulties for Mohammed and other consumers like him.

The first factor that could impact consumers' comprehension is script and schema activation. Specifically, scripts are time – ordered schemas. For example, going into a new restaurant, a person knows the time order in which different behaviors should occur. A person first sits down and has water arrive, then is given time to look over a menu and decide what to eat. These behaviors and interactions with staff follow the script of going out to eat. The restaurant script aids in understanding of the situation. With this in mind, the idea underlying scripts is that when people read text, time-ordered data structures called scripts are activated to aid understanding (Abelson, 1981). Scripts can lessen the cognitive burden of processing a stimulus since they allow a person to rely on the previously stored information or representation of that stimulus and are present during encoding to help select and organize the information that is relevant into memory. They can be described as a library for associations because incoming stimuli and information will activate associated information that is related to the incoming information. The activation of scripts is automatic and unable to be turned off, which is usually necessary

and useful, but can sometimes contribute bias or inaccurate interpretations (Smith & Houston, 1985, p. 214).

In addition to the specific time – order of a script, schemas are thought of as mental representations of the world and objects but can be applied to cognition as a structure for text comprehension. Schemas function as prediction mechanisms for ambiguous information or information that is not fully understood (McVee, Dunsmore, & Gavelek, 2005). For example, if someone asks you to guess what animal four legs has, nodes that contain the mental representation and script of animals that have four legs will begin to activate. If you were told the animal has four legs and is also a household pet, the activated script would shift and may allow you to guess the correct animal since the possibilities are smaller to pick from. If it was added that the animal likes to chase frisbees and bark, you would instantly be able to guess the animal. In the context of text comprehension, a reader goes into a text with their expectation of what the text is going to be about, as well as their prior knowledge about the concepts presented. For example, in a study by Bower el al., (1979) subjects were more likely to remember and recall stories that had action events closer to where a participant expected them to be in a sentence. When presented with a text, the reader uses their organized schema structure to link what is read in the text to what they already understand about the text. This is a way to combine individual data structures representing different concepts into a whole overall idea (McVee, 2005). These inferences may not always be correct, depending on what the text is. The function of this schema is to allow for faster and more automatic processing of a stimulus, but at the expense of accuracy (Smith & Houston, 1985, p.215). For example, when reading a difficult to read text, the reader is unable to activate any type of

script or schema to base inferences on since they would not fully comprehend what was read. This would lead to gaps in their understanding of that text. Without script or schema activation, the reader is unable to create a "hypothesized cognitive structure" which does not allow them to make inferences and fill in gaps when they lack understanding (Abelson, 1981).

The actual and semantic meaning of a sentence is not stored, but the idea or abstract thought that is communicated through the sentence is what gets stored in a schema. Once stored, the packets of information are connected in larger networks that allow for more generalizations and predictions related to a script that becomes active (Abbott, Black & Smith, 1985, p.179). As a reader, a schema aids in "planful retrieval" and allows the reader to reconstruct meaning (McVee, 2005, p. 537). When Mohammed was faced with difficult text in the loan, he may not have had the schema activation needed to help him figure out and comprehend what he was reading, since he was most likely not familiar with the language in legal type documents and did not have enough prior experience or exposure reading that type of legal material.

From research on scripts and schemas, it is clear the knowledge and expectations of the reader impact comprehension of a text. Script research relates to a theory proposed by Weaver and Bryant (1995) that talks about optimum effort, which is the idea that readers can comprehend texts that are matched to their reading level, instead of above or below it. In passages that are too easy, the reader may skim through the text in an "automatic role" and not truly comprehend what was read when they were asked about it later. The readers may assume that since the text was easier to read, it would be easier to understand it (Weaver & Bryant, 1995, p.12). In a more difficult text, the reader would

exert much more cognitive effort on trying to read through the material than on understanding what they read at the end of the text. A harder text can deplete cognitive resources, and therefore, the reader may not be able to fully synthesize the material. Attention, memory, and visual processing would all be at risk if cognitive resources became too depleted. If a reader cannot fully comprehend a text, they would have to revert to filling gaps in their comprehension with their scripts, which might not be fully accurate (Weaver & Bryant, 1995, p.17). While reading an easy to read text, the reader can represent larger packets of information about that text in a smaller amount or single element in working memory, which would allow them to experience less cognitive load and difficulty when trying to encode and decode the information they have read (Sweller, 2011).

A follow up study conducted by Lin et al., (2002) tested the theory of optimum effort hypothesis and found that even if the grade level of a text was a match to the participant's level, the readers still had trouble comprehending a difficult passage of the same grade level. Results of that study found that for each reader, there is an optimum readability or grade level of reading that would most likely be understood, with an average level of cognitive effort. If a text is written at the same readability level as the reader, then that reader does not have to use excess cognitive resources to understand and will have a better opportunity to comprehend what was read. A text that is harder to read would require more effort, while a text that is too easy to read could become boring and cause distraction since no effort would be required. If the text was too difficult, the reader would not understand the sentence at word level, and they would not understand how the words are related to the overall sentence (Lin et al., 2002, p.196). In this sense,

characteristics of the reader play a role in comprehension of text, with general knowledge and motivation to read included as important factors. If a reader is more motivated to read the text, they would put more effort into comprehending what they read, rather than if they were not motivated and simply skim through it.

Aside from scripts, working memory abilities and individual differences may be another factor that contributes to comprehension differences. Working memory provides for the temporary storage of information and integrates relevant information with long term memory, while inhibiting irrelevant information from being processed (Elsayyad, Everatt, Mortimore, & Haynes, 2017). The capacity of working memory is not unlimited and can only store a certain amount of information, which was generally thought to be seven plus or minus two items (Miller, 1956). However, this number has been debated in recent years, with some studies claiming it is too high (Cherry & LeCompte, 1999; Cowan, 2010). For example, a study by Cowan (2010) found that for adults, the immediate memory span capacity was only around three to five items or chunks. Although each chunk can contain more than one item, they would still be held in working memory as one entity which is a way to increase working memory capacity. Applying this to the current study, the participants were not legal experts. This meant that participants in the difficult to read condition would have a much harder time holding the information in their working memory because the passage would be unlike anything they would have read before and therefore harder to create chunks of similar information to store in working memory. The difficult to read condition would put more strain on working memory and potentially use more mental resources than the easier to read passages. Although there are individual differences in working memory capacity, it plays

the role of holding current and relevant information easily accessible in order to be used or applied. In the case of reading contracts, working memory is attempting to hold the relevant information in chunks, if possible, in order to give the reader the best possible interpretation of the material in the least amount of cognitive space needed. If a contract was particularly difficult and hard to read, working memory would struggle to process the information and would not be able to fully aid in comprehension. This could have been a large factor that caused the problems for Mohammed, since legal writing normally exceeds almost every non-lawyer's working memory capacity.

In research by Rumelhart and Ortony, they mention that readers may simply skip the unfamiliar words and just reconstruct a partial explanation when they are tested on their comprehension (1997). The reader may exert minimum effort to understand what was read, and since trying to decipher a more difficult word would require more cognitive effort, the reader may choose to not exert that effort and try to piece together comprehension of the text at the end (Rumelhart & Ortony, 1977, p.111-115). The rare or more difficult words in a text "signal" the reader that what they are reading may require more cognitive effort and therefore contribute to the reader not fully reading what they are presented or just skimming for main themes. This type of signal may prompt readers to skip that section rather than try to work through it. For example, in the document that Mohammed signed, there may have been very rare and difficult words that he, or many other people, may not have never been exposed to. This would be a word that is too difficult to dwell on to try to understand which would result in skipping that sentence or maybe even the entire paragraph.

Research by Cherry and LeCompte (1999) tested the idea of mental strain on working memory and found that the length and meaning of the word contribute to more strain on working memory capacity (p. 289). The strain would come from the inability to actively chunk familiar words and concepts together in order to attempt to increase working memory capacity. The low readability text is much harder to read because there are longer and less familiar words. Even if a reader were able to parse, with effect, what each of the individual words meant, this would take more working memory resources to comprehend the meanings of sentences, paragraphs, and passages. These working memory resources might then be exhausted when they try to comprehend the subsequent passage. Words that are more familiar are remembered better than words that are unfamiliar and do not hold as much meaning. For example, if a participant tried to activate their schema while reading the difficult to read passages, they would not be able to store the same amount of text information in the same amount of working memory space as the easier to read text. The harder text would not be able to activate a schema in the same way as the easy text because participants would not have as much prior information or knowledge to activate the script in a useful way. If a schema is not activated, the reader would not be thinking in the correct way about the text and could lead to less focus and attention on the material to read. On top of this, if working memory capacity is less, or if chunking is not readily available due to the unfamiliar language, then readers who are tasked with reading a difficult text would have a much harder time chunking the information and processing it in their working memory. Individuals with low working memory capacity may also be more prone to shifts in attention and are unable to keep retrieve and filter information related to their overall goal (Unsworth &

McMillian, 2013, p.832). In the case of Mohammed, he may not have been exposed to the specific loan contract language enough to be able to form and process active chunks of information that would have aided his working memory capacity and overall comprehension.

Attention has also been an important research topic in reading comprehension literature and could be part of an explanation why readers may struggle to comprehend a subsequent passage after reading a difficult to read text. Working memory function can be impacted by mind wandering, which is the idea that a reader's attention gets shifted to focus on internal thoughts and feelings from the external information (Unsworth & McMillian, 2013, p. 833). In other words, when a participant is not paying attention or if their minds are wandering, then they will not understand the text. Shifting cognitive attention and resources would make it more likely for the reader to miss something in the text or not focus on what they are reading.

Mind wandering, and attentional shifts have been found to be related to concepts such as motivation and interest (Unsworth & McMillian, 2013, p. 834). For example, if a reader does not have high motivation or interest to focus on the text, then their attention and cognitive focus would not be fully on the material and contribute to mindless reading. Mindless reading is simply skimming a text just looking for potential keywords that would sum up what you read. However, when the text is difficult to read, a reader would have a very hard time skimming the text since most of the words and phrases would not activate an available script for the reader, which would not give the reader any way to fill in gaps in their comprehension with previous knowledge or correctly guess what they read. A previous study on mind wandering found that thoughts that were

unrelated to the text were negatively correlated with accuracy on a comprehension test after reading a passage (Schooler, Reichle, & Halpern, 2004). As a reader allowed themselves to think and focus on other things besides the text, they were more likely to not fully understand what they had read. If participants are focused on other things and not keeping their working memory resources focused on the reading task, then they will not be able to fully comprehend that material. People experience mind wandering every day, especially in situations where there are many distractions, or the task does not meet a certain level of arousal. For Mohammed, there may have been distractions, internal or external, present while he was trying to go through the contract which could have created problems.

Most of the current research in this topic has focused on factors that contribute to people reading in the first place or are focused on the language side of comprehension. Stolle and Slain (1997) found that when they asked participants to rate how likely they would be to read a contract, the average response was "somewhat likely, but would be less likely to read the contract closely," (p. 93). These researchers also found that only two-thirds of the participants understood what the clause they read was about. Before even taking text difficulty or readability levels into account, the participants from this study responded as if they would not read a contract closely and potentially just gloss over it. A study by Stark, Choplin, and Linnabery (2013) attempted a similar replication, but applied the results to a loan contract. They tested participant's ability to understand what they were reading by wording three clauses with language that contained either fair, clearly unfair, or vaguely unfair clauses. They found that fifty-five percent of participants failed to understand the impact of the clause wording on their rights, which "reflected a

lack of understanding of the impact of the contract language" (p. 810). If consumers are less likely to even read clauses in the first place, their ability to understand them would be drastically decreased before they would even start reading.

What happened with Mohammed Hoque and many others raise questions as to how loan contracts can be presented in different ways that may increase or decrease reader comprehension. The previous studies have shown a gap in research that has not covered how text difficulty order is an impactful feature in texts or contracts. If the language in contracts considers the effect that a preceding difficult clause may have on what follows it, then this could help to improve the overall comprehension of those who sign them. The current research sets to explore the relationship between prior text readability and overall comprehension of a subsequent text in the context of loan clause language. This study will explore how an easy or difficult to read text can impact the subsequent passage and overall comprehension. For the purposes of this proposal, readability is defined as the level of complexity of a text relating to the average grade level of reading (Flesch, 1948). Reading comprehension is defined as the act of a person understanding what a text is communicating and is measured with sentence verification questions (Marcotte & Hintze, 2009).

Rationale

A related phenomenon in the current body of research is the spillover effect, that studies how the fixation time on words can be impacted by the previous word. (Shvartsman, Lewis, and Singh, 2014). This proposal is different from the spillover effect research because the text as a whole is used as the preceding variable rather than just a word. Aside from the spillover effect, the current research body has focused more on how

people read and process that information, not how the readability of a preceding text can influence comprehension of the subsequent text. This proposal attempts to fill a gap in research by manipulating factors in the first presented text that may increase or decrease comprehension of the subsequent one. Additionally, other studies that have conducted readability research do so in the context of education of students and teaching students to read, rather than applying it outside the classroom. These studies also seek to apply the concepts and findings to other realms, specifically contracts and written agreements in law settings. Previous research has shown that people do not always closely read texts that they are given, which contributes to a decrease in overall understanding. The goal of this proposal is to demonstrate how readability levels of a presented text can impact comprehension of the subsequent text, which should be considered by both the consumer and the creator of the document. Making changes to the structure of documents can increase the flow of communication from all parties involved.

Pilot Study

The pilot study explored the influence that an easier to read or difficult to read preceding clause had on comprehension of the subsequent clause. The target subsequent paragraph was a trial by jury clause that the comprehension questions were based on. The target paragraph was the subsequent clause in both the easier and difficult to read conditions, but was the only paragraph presented in the control condition. The demographic variables of age, gender, and loan experience were collected at the end of this study as well. The goal of the pilot was to identify an effect of readability on comprehension of the subsequent paragraph.

Pilot Study Hypothesis

Participants who read a low readability (harder to read) paragraph first will score lower on comprehension questions about the subsequent paragraph than participants who read a higher readability (easier to read) paragraph first. The subsequent paragraph, also called the target clause, was a trial by jury clause that the comprehension questions were based on. The target paragraph was presented after the easy and hard paragraphs in those conditions, but was the only paragraph presented in the control condition. The control condition did not have a preceding clause before the target clause.

Method

Participants

Participants were recruited through MTurk and given ten cents for their participation in the study. A total of 149 participants responded to the survey, but data from nine participants was excluded for completing the survey in under one minute. Data from one more participant was also excluded because they did not answer any comprehension questions. After exclusions, fifty-five participants identified as male, eighty-two as female, and two did not answer. The ages of participants ranged from twenty to seventy-seven, with the median age being thirty-five, while the average age was 37.5. In this study, ninety-nine percent of participants marked that they had taken out a loan before.

Materials

This pilot study was designed to test whether a preceding clause's readability level has an effect on comprehension of the subsequent text. For example, if an extremely difficult to read clause is presented, participants may struggle to read that clause since it would take up working memory and other cognitive resources. If they struggled for

longer, they might just skip over it, inadvertently skipping over the target clause that is presented after it. Therefore, they would not be able to answer the comprehension questions correctly. This study used a clause that was pulled from a loan document. The clauses chosen were not written in any specific layout or universal format, so if someone had experience with contracts they would not have an advantage reading.

Three versions were manipulated to create three different readability conditions. The first condition contained prior text that was extremely difficult to read (low readability), the second contained prior text that was very easy to read (high readability), and the final condition was used as a control condition, which had no preceding text (See Appendix A*)*. To hold other variables constant, the number of words were balanced across each passage and the clauses were used from similar loan documents. In addition, the meaning of the words in the easier and harder to read versions were as close as possible, since synonyms were used. How these manipulations affected readability was measured using the Flesch-Kincaid readability index.

The Flesch-Kincaid readability index (Kincaid, et al., 1975) uses average sentence length and average number of syllables per word to calculate a readability score, with higher scores translate to lower levels of readability. The design of this pilot study randomly assigned participants to one of three groups. One group was simply used as the control group, which was only given the target paragraph. This was the paragraph the comprehension questions were all based on and did not contain common knowledge. This section had a readability score of 13.5 and was not included in scoring the readability of the difficult or easy to read paragraphs. The target paragraph was the only clause presented for participants to read in the control condition. The other two conditions were

the hard to read and easy to read paragraphs. In both of these conditions, participants either read the easy to read or hard to read clause first and then were presented with the target section to read after and then answer the comprehension questions. The difficult to read condition had a score of 29.1 on the Flesch-Kincaid readability index. This score did not include the attached target clause, the readability is only of the very difficult to read clause. The easier to read condition (not including the target clause) had a score of 13. The only difference between the three conditions was what the preceding text (if any) that came before the target passage. The target clause was the same in all conditions and was sufficient for participants to answer the questions used as our dependent measures correctly. The target clause was only four lines and created to be between the easier and harder clauses. Due to how the target clause was created, it was closer to the easier to read paragraph readability than the harder to read one. However, this clause was chosen since it was not too easy or difficult to read and was straightforward enough for participants to answer the comprehension questions about it with ease, if they read it (see Appendix A*).* All the comprehension questions were based on the target clause. Also included in the easier and difficult to read paragraphs was the sentence, "The glow of the moon illuminated the lake." While the control paragraph had the line, "The cheese was not moldy." This was done to be a check to see if participants read at all. The creation of this measure is further explained below.

Comprehension Questions

The primary dependent measure was a series of five hypothetical scenario "yes" or "no" questions used to measure comprehension of the target trial by jury clause. These hypothetical sentences paraphrased part of the trial by jury portion and were based on

what was read.   An example of the hypothetical comprehension scenario was: "You had

a falling out with your business partner and they now want to back out of the loan

contract they signed with you and pull their money. Knowing that your partner was

supplying half the investment, you realize that you can't afford the loan amount that was

originally provided, so you want to talk to the lender to get a smaller loan amount. The

lender refuses to reduce their investment, and you want to take them to court to get the

loan reduced since you lost one of the signers. Are you able to have a trial by jury?" (See

Appendix A for full list). These hypothetical scenarios were created to contain

information about the trial by jury waiver and include a situation in which someone wants

to hold a trial by jury and if they were able to, based on the information given in the

scenario. Participants were given either a yes or no bubble to click, so they did not

answer with any other written comments. For a single participant, scores ranged from

zero correct to five correct on this hypothetical comprehension measure. A score of five

would mean that participant answered all the hypothetical questions about the trial by

jury clause correctly. Scores were calculated by taking the number of correct answers out

of five. The comprehension questions would not be able to be guessed by anyone who

was familiar with contracts either, since they were novel scenarios.

Reading Checks

A second dependent measure looked at whether participants read at all. This

measure was used as a reading check and included two additional questions that asked:

"Did you read a line about the moon?" and "Did you read a line about cheese?" These

questions were embedded in the middle of the paragraphs and were made on an unrelated

topic so if the participants did read the clause, they would remember seeing those lines.

The question that asked this measure was just scored out of two by checking if the participant said "yes" or "no" to whether they had seen and read each line in their paragraph. After running the pilot study, the reading checks were updated to not be a hit-only criterion and added two more questions. Two more reading checks were added that required an answer of "no" to be correct. This was done to bring the correct answers of all four reading checks to have two "yes" and two "no." These changes created more reliable reading check answers, rather than having them all be correct with the same answer. The two new reading checks were applied to study one and two, which asked if participants read a line about a cow or if they read a line about an alien, both of which are not mentioned in any paragraph, so the correct answer was "no." This was done to address issues of guessing and make the reading checks a more viable measure. The reading checks were updated for the analysis of studies two and three, but not for the pilot study since that study was already conducted.

Procedure

The study was posted on MTurk and was advertised to the participants as a test of puzzle-solving skill and reaction time first and finishing with puzzles offered at the end of the study. This study was advertised as a puzzle task in order to put participants in more of a real word situation where they were reading through the contract in order to get to the end goal of the puzzle. People sign contracts in order to obtain some type of end goal and the puzzle task was meant to be something desirable for the participant to work towards after reading. The actual informed consent was presented first, but also included the loan paragraphs that were manipulated so they just seem like a longer section of the consent form. The formal language in the consent is similar to the language used in the

actual readability conditions, so it blends in and looks more like a coherent structure. Participants were randomly assigned one of the three versions. If they were in the high readability condition, they were given the easier to read clause followed by the trial by jury paragraph and if they were in the low readability condition, they were given the harder to read clause followed by the trial by jury paragraph. Participants randomly assigned to the control group were only presented with the single trial by jury target paragraph to read. After reading through their paragraphs, participants answered the five hypothetical comprehension questions based on the target clause and then the two reading check measures.

Results and Discussion

Prior to the analysis, data from participants who finished the survey in under one minute were excluded, because it was not possible to read everything that was presented in one minute or less. Nine participants had their data excluded.

Reading Checks

The reading checks were analyzed first to see if participants read the materials close enough to observe them. A participant was classified as failing to read if they gave an incorrect answer to one or both reading check questions. In the control condition, to get both answers correct participants had to correctly say that they did see a sentence about cheese being moldy and also correctly say that they did not see a sentence about the glow of the moon. In the two experimental conditions (the easy condition and the difficult condition), to get both answers correct participants had to say that they saw a sentence about the moon and the cheese. The checks were sentences that were embedded in the readability paragraphs and were scored on a scale from zero to two.

A one- way ANOVA was conducted to compare the effect of preceding paragraph readability level on subsequent paragraph reading check scores. The ANOVA did not find a statistically significant difference in reading check scores based on readability group, $F(2, 135) = 1.22$, $p = 0.299$. Table one shows the marginal mean reading check score for each condition.

Condition	Marginal Mean	Standard Deviation
Hard	0.37	0.75
Easy	0.24	0.58
Control	0.17	0.53

Table One: Marginal mean reading check score by condition.

A potential floor effect may have occurred if the clauses were too long for participants to give their full attention, or the reading checks were embedded in a spot that participants were more likely to skip over. This reading check measure was only based on two questions and was updated in the following studies to be a more reliable measure.

Comprehension Questions

For this pilot study, comprehension questions were scored out of five and this score was the main dependent measure. After each participant's scores were totaled, a one- way ANOVA was conducted to look for a significant effect of preceding paragraph readability level. The one-way ANOVA revealed a statistically significant difference between readability condition and comprehension score, $F(2, 135) = 6.367$, $p < .05$. A Tukey follow up test found the difference was between the control ($M = 3.781$, $SD = 1.41$) and difficult to read group ($M = 3.18$, $SD = 1.31$); $p < .05$. Consistent with predictions, participants in the control condition answered significantly more

comprehension questions right than did participants in the difficult to read group. However, there was not a significant difference in comprehension scores between the difficult to read group and easier to read group ($M = 3.12$, $SD = 1.19$), $p = 0.528$. The easy to read condition was also not significantly different from the control condition, $p = 0.059$. See table two for the comprehension score marginal means.

Condition	Marginal Mean	Standard Deviation
Hard	2.82	1.10
Easy	3.12	1.19
Control	3.71	1.41

Table Two: Marginal mean comprehension scores at each readability condition.

Although the predicted difference in comprehension scores between the hard and easy conditions was not statistically significant, this pattern of results is generally consistent with the view that text difficulty can create extra challenges for participants as they try to read and comprehend subsequent text.

The pilot study found a significant effect of readability condition. The comprehension scores for the control condition were significantly different from those of the low readability condition. The participants in the control condition had higher comprehension scores than participants in the harder to read condition. Although there was not a significant difference between the hard and easy to read conditions, the pilot provided insight into this phenomenon.

Power Analysis

After the pilot study was conducted, an effect size was calculated using Cohen's *d*. This compared the effect size between the high and low readability conditions. The

low readability condition (n = 38) had a mean of 3.18 and a standard deviation of 1.31

and the mean of the high readability condition (n = 41) had a mean of 3.37 and a standard

deviation of 1.32. These data were used to calculate a $d = .144$. This effect size is

obviously a very small effect size, which is further addressed in the limitation section.

This effect size was then used to calculate an A priori power analysis prior to running the

next two studies to figure out how large of sample and group sizes would be needed to

get power of 0.8 based on the effect size that was calculated with the data from the pilot

study. This analysis showed that these studies would need around 758 participants per

condition and a total sample size of 1,516 in order to have the power of .8 needed to

detect the small effect size on comprehension scores. This is also addressed in the

limitation sections of this paper.

Smallest Meaningful Effect Size

A second power analysis was conducted to find the number of participants

required to detect an effect if the effect size was set at the smallest meaningful effect size.

Since this study focuses on research that has not been thoroughly tested, there was not

any previous power analysis data to define a meaningful effect size. Instead, the smallest

meaningful effect size was set at .3 because it is regarded as a small effect size in other

literature. The power analysis was conducted with .3 as the effect size and calculated the

number of participants needed in each condition in order to detect an effect that size with

power of 0.8. The results of this power analysis showed that each condition would need

176 participants in order to detect the smallest meaningful effect size of .3.

The results from the pilot study led to the creation of a second study that focused

on how the length of the preceding clause may impact comprehension, which was one of

two independent variables explored in the following study. Although these results found a significant effect, there were problems with low power and small sample sizes, which are discussed further in the limitations section. The results of the pilot demonstrated that there is a potential effect present and allowed further exploration of this topic and provided reason to conduct follow-up studies.

To further investigate this phenomenon, two follow-up studies were conducted. The first study investigated the effect of paragraph length and readability by creating shorter and longer versions of the original paragraphs. The second investigated the impact that working memory has on the ability to read and comprehend the loan paragraphs.

Study 1

The final two experiments were pre-registered on OSF (open science foundation) in order to avoid bias and type one error prior to observation of the data and data analysis. Changes to the analysis plan were also added as updates to the pre-registration prior to actually running any analysis. Updates after a review were added to reflect the changes made to scoring.

This study was developed from the results of the pilot study and used the same materials and procedure, except it was conducted in a 2 (readability) x 2 (text length) design. The first independent of readability consisted of the easy to read and more difficult to read conditions from the pilot study. The second independent variable of length consisted of shorter versions of both the easier and difficult to read clauses that were used in the pilot study. This condition was made up of a short easy to read condition and short difficult to read condition. The goal of this study was to rule out a possible

confound of paragraph length from the pilot study, since the control condition was the shortest condition and had the highest mean comprehension score. Participants answered the same comprehension questions after reading the preceding and subsequent paragraphs.

Hypothesis I

Participants in the short length condition will score higher on the comprehension questions across all levels of readability, with the highest comprehension scores coming from the easier to read, short paragraph length condition. Replicating the pilot study, I predict a main effect of readability as well as a main effect of length. I would not expect one effect to be super-additive in a condition, so an interaction is not expected. Participants will score higher on the comprehension questions if they are in the easy condition compared to the hard condition as well as score higher in the shorter condition compared to the longer condition.

Method

Participants

Participants were recruited through DePaul University's undergraduate subject pool SONA system and given 1/2 credit for their participation. Data from 117 participants was collected, but ten were immediately excluded because they either finished the survey in under two minutes or they did not answer all seven comprehension questions. Additionally, six more participants were excluded because they did not get at least two of the four reading checks correct. Of the 101 participants included, seventy-two identified as female, twenty-two identified as male, three identified as nonbinary, and

four preferred not to report. The age range of the participants was from eighteen to forty-four, with a mean age of 19.9 years.

Materials

The four conditions were created, with each participant assigned to one of them for the 2 (Readability: easy or difficult) × 2 (Text Length: short or long) between-subjects design. The two long conditions in this study consisted of the original hard to read and easy to read paragraphs that were used in the pilot study. The longer easy to read condition had a readability score of 13 and a word count of 297, while the longer hard to read condition had a readability score of 29.1 and a word count of 299. To make them shorter, longer words were replaced with shorter synonyms and information that was irrelevant to understand the comprehension questions was deleted. This was done to make the shorter and longer versions of the same clauses read as similar as possible. The shorter versions had word counts approximately half the size of the longer paragraph conditions.

The new shorter and easier to read paragraph had a readability level of 12.6; word count of 151, while the short and more difficult to read condition had a readability level of 21.4; word count of 152. The shorter harder to read clause was not as difficult as the longer version, which makes comparisons more difficult. However, the readability level is still drastically more difficult than the easier versions. The target clause was also changed to become shorter and had a new word count of 75 and was still the subsequent clause in the four conditions. This study did not have a separate control condition with only the target clause. The same demographic variables of age, gender, and loan experience were asked at the end of this study.

Comprehension Questions

The same hypothetical scenario comprehension questions that were used in the pilot study were included in this study. The comprehension questions only refer to the target paragraph, which is the same in all of the four conditions. An example of the hypothetical comprehension scenario was: "You had a falling out with your business partner and they now want to back out of the loan contract they signed with you and pull their money. Knowing that your partner was supplying half the investment, you realize that you can't afford the loan amount that was originally provided, so you want to talk to the lender to get a smaller loan amount. The lender refuses to reduce their investment, and you want to take them to court to get the loan reduced since you lost one of the signers. Are you able to have a trial by jury?" (See Appendix A for full list). For a single participant, scores ranged from zero to five for the correct answers on this hypothetical comprehension measure. A score of five would mean that participants answered all the hypothetical questions about the target trial by jury clause correctly. Each condition answered the same comprehension questions which included three questions that would be correct with an answer of "yes" and two questions that would be correct with an answer of "no." This was done to increase the reliability of the comprehension questions and not allow participants to get high scores by just answering all yes's or no's.

Reading Checks

After feedback from the pilot study, the reading checks for the final two studies were updated. Instead of just two reading checks, four questions were presented in each of the four conditions to assess reading. Each condition included the same four reading checks, two of which needed to be answered with a "no" to be correct, and two that were

required to be answered with a "yes" in order to count as correct. In total, the reading checks asked participants if they read a line about the moon, cheese, a cow, or an alien. The cheese and moon questions were embedded in the text, but the questions that asked about a cow and an alien were not included in the text of any conditions. Since each condition had the same number of reading checks that had to be answered the same way, each condition was scored from zero to four. A zero meant the participant did not get any of the reading checks correct and either did not see any of the checks when they were present or said they read them when they were not present. A score of four meant that they read them when both checks were present and marked that they saw them or marked they did not read them when the final two checks were not present. As mentioned earlier, a participant had to answer at least two out of the four reading checks correct to have their responses included in analysis.

Procedure

Rather than describing that this study was about reading comprehension and loan clauses, participants were told that this survey was about puzzles and would have a chance to attempt puzzles after completion. This was done to try to add more intrinsic value to the survey, so participants may have wanted to try harder if it was about something that was more interesting to them than loan clauses.

A 2 (Readability: easy or difficult) × 2 (Text Length: short or long) between-participant design was created. Participants were randomly assigned to one of the four conditions and answered the five hypothetical situation comprehension questions and four reading check questions.

Comprehension Questions

The comprehension questions were the same scenarios used in the pilot study. For a single participant, scores ranged from zero to five for the correct answers on this hypothetical comprehension measure. Each condition answered the same comprehension questions which included three questions that would be correct with an answer of "yes" and two questions that would be correct with an answer of "no." This was done to increase the reliability of the comprehension questions and not allow participants to get high scores with just answering all yes's or no's. The score from the five comprehension questions were added to the score from three of the reading checks, for a final score out of eight.

Reading Checks

For the reading check questions, only three of the four were used in the final total scoring. The one reading check that was excluded from the final total score was the only reading check that was present in the preceding paragraph and was not included because this study was focused on comprehension of the subsequent paragraph, not the preceding, The reading checks that were used in the final total score were the two that were not present in the material that participants needed to answer with a "no" and the one reading check that was actually present in only the target clause. In other words, the three reading checks included asked about cheese, a cow, and an alien. The question about cheese was included in the subsequent paragraph, while the cow and alien questions were not mentioned in any condition and required a "no" answer to be correct. The two questions that were not embedded in either paragraph were included in the total score in order to increase the reliability of the overall measure. In other words, those three reading checks

were chosen because they included questions about the subsequent paragraph that was not directly included in the preceding paragraph. In total, the comprehension measure was scored out of eight since the five comprehension questions were combined with the three reading checks. The same demographic variable questions of age, gender, loan experience, and education level were also recorded.

Results and Discussion

A 2 (Text Length: short vs hard) x 2 (Readability: hard vs easy) factorial ANOVA was conducted to analyze the effect of length and readability on comprehension question scores. The two – way ANOVA revealed there was not a statistically significant interaction between the effects of length and readability on comprehension scores, F $(1,97) = .0.494$, $p = .484$. Table three shows the marginal mean comprehension score for each specific readability by length condition.

Condition	Marginal Mean	Standard Deviation
Hard and Short	4.22	1.273
Hard and Long	4.52	1.27
Easy and Short	4.17	1.271
Easy and Long	4.12	1.269

Table three: Marginal mean comprehension scores at each interaction condition.

Main Effect of Readability	Marginal Mean	Standard Deviation
Hard Readability	4.37	1.29
Easy Readability	4.14	1.27

Table four: Marginal mean comprehension scores at each readability level.

Main effects analysis showed that paragraph length did not have a statistically significant effect on comprehension scores, F $(1,97) = .223$, $p = .638$. There was not a significant difference in total comprehension scores based on whether participants read a

longer (*M* = 4.32, *SD* = 1.271) or shorter (*M*= 4.20, *SD* = 1.273) paragraph. In addition,

the main effect of paragraph readability did not have a statistically significant effect on

comprehension scores either, *F* (1, 97) = 0.798, *p* = .374. There was not a significant

difference in comprehension scores based on whether a participant read a hard to read

clause (*M*= 4.37, *SD* = 1.29) or an easier to read clause (*M* = 4.14, *SD* = 1.27). Table four

contains the marginal means for each main effect.

Main Effect of Length	Marginal Mean	Standard Deviation
Long	4.32	1.271
Short	4.2	1.273

Table five: Marginal mean comprehension scores at each main effect level.

The goal of study one was to test for differences in comprehension score based on

whether a participant read a long or short paragraph first, or an easy or difficult to read

paragraph first. The original difficult to read and easy to read loans clauses from the pilot

study were used as two conditions, and two new conditions were created that were

shorter versions of the original readability paragraphs.

The results did not find a significant main effect of length on comprehension

scores or a significant main effect of readability level on comprehension scores. The

interaction between length and readability level was also not significant. Although the

predicted results were not statistically significant, each condition had less than twenty-

five participants, so this study may have been too low powered to detect any significant

effects. Additionally, the readability level of the short and low readability condition was

not as difficult as the long low readability paragraph, which may have impacted the

overall difficulty of the short paragraph and made that condition easier than it was intended. The mean comprehension scores are all slightly above chance and could be attributed to how much participants engaged with the material or if the comprehension questions were too difficult for participants to answer. Study two is designed to test readability and comprehension but adds a new measure that explores how verbal working memory impacts the process.

Study 2

The final study examined how working memory capacity affected participants' abilities to read and comprehend the subsequent target clause. The previous studies focused on factors that can impact comprehension of the subsequent text by manipulating the readability and length of the prior text, while this study investigated how individual differences in working memory may play a role in their ability to comprehend the target jury waiver clause. A working memory measure adopted from Daneman and Carpenter's (1980) was used to score the memory portion of this study and will be described in the materials section. An ANCOVA analyzed the data and included total words remembered as the covariate to examine the differences in comprehension scores between each condition after controlling for verbal working memory. The dependent measure was the same measure used in the previous study and included the score out of five from the comprehension questions, added to a score out of three from three of the four reading. checks. This study was developed to explore the hypothesis that participants with higher working memory capacity would score higher on the comprehension scores. Participant differences were not examined closely in the previous two studies and could have been a

factor in why one participant may be able to correctly answer more comprehension questions than another, regardless of what readability level they have to read through.

Hypothesis II

The low readability (difficult to read) passage would consume more cognitive resources since the participants would have to work harder to read, with less help from their current scripts. Therefore, the difficult to read passages would decrease the ability to activate and rely on a known script for shorter processing and comprehension of the text. Participants with higher working memory capacity were predicted to comprehend the subsequent target clause better than participants with low working memory capacity when reading the difficult to read clause first. It was hypothesized that participants with high working memory capacity would be able to overcome the barriers created by the difficult to read, low readability clauses.

In addition, the differences between the high and low working memory capacity groups will be much smaller for participants who read through the easier to read loan clause first, because the easier clause will present a much smaller barrier to comprehension for the low memory capacity participants to overcome. Testing the impact that working memory capacity has on the ability to hold and process more difficult information offers insight into how a reader may work through or struggle while reading the difficult text, and how comprehension of the subsequent paragraph would be affected.

Method

Participants

Participants were recruited through DePaul University's undergraduate subject pool SONA system and given 1/2 credit for their participation. A total of 102 participants

signed up for the study, however data was excluded from sixteen immediately because they did not fully complete the survey or submitted it in under five minutes. Additionally, data was excluded from six more participants for failing to score at least sixty percent on the sentence verification questions. This measure asked participants after each sentence if that sentence was a correct sounding English sentence. A participant had to correctly answer at least sixty percent of those questions correctly to have their data included. The last exclusion criteria focused on the reading check questions. A participant was excluded if they did not score at least two out of four on the reading check questions, which led to the exclusion of data from five more participants. After all of the exclusion criteria, the total number of participants included in the analysis were seventy- five. Of the participants included, fifty- one identified as female, five identified as non-binary, and nineteen identified as male. The range of ages was between 18 and 28 while the mean age was 19.6.

Materials and Procedure

This study began with participants randomly assigned to either the easy to read, difficult to read, or control condition. The three conditions were the same three used in the original pilot study. The difficult to read condition (low readability) again had a readability score of 29.1, the easy to read condition (high readability) had a readability score of 13, while the control condition, which just presented the target clause to participants, had a readability score of 13.5. The difficult and easy conditions both had the target passage presented immediately after them, while the control condition only had participants read the target clause without anything before it. The same comprehension question and reading check measures from all three studies were also used again. The

only difference in the final study was an added measure of verbal working memory at the end. Participants read the paragraph based on their condition prior to the target clause and then answered the same five comprehension questions and the same four reading check questions. After this, they were given the verbal working memory portion of the study, which will be described in further detail in the section below.

Comprehension Questions

After reading the paragraphs, participants answered the same five comprehension questions, three of which would need to be answered with a "yes" to be correct and the other two need to be answered with "no" in order to be counted as a correct answer. These comprehension questions were the same in all conditions and only asked about the subsequent target clause. A score of five would mean that participants answered all the hypothetical questions about the target clause correctly. This score out of five was added to a score out of three that came from three of the four reading checks. The total score for the dependent measure was out of eight.

Reading Checks

For the reading checks, each condition had the same four reading checks as well as the same sequence of correct answers. Two reading checks required a "yes" to be counted as correct, and the other two checks required a "no" in order to be correct. The same reading checks that were used in the previous study were used again in this study. They asked participants if they read a line about the moon, a line about cheese, a line about a cow, or a line about an alien. Since the line about cheese was in all of the conditions' target clauses, that reading check was included in the final dependent measure. In addition, the two reading checks that were not included in any paragraph and

required an answer of "no" to be correct were also included in the final measure as well. As with the previous study, scores from the reading checks that asked about cheese, a cow, and an alien were added to the scores from the five comprehension questions. The reading check that was included in the preceding paragraph was not counted since it did not relate to the target clause. A participant's data was excluded if they did not answer two out of the four reading check questions correctly.

Working Memory Test

After they completed the comprehension questions about what they read, they were given a working memory capacity test, which was the verbal working memory span test based on Daneman and Carpenter (1980). For this test, participants were presented with groups of two to six sentences. After reading each sentence, they had to identify whether or not that sentence was a correct sounding English sentence. This sentence identification task served as a grammatical judgment for the participants. The task asked participants if the sentence they read sounds correct and makes "sense" in English or not. An example of a sentence that would have received a "no" answer is, "After exams all the passing, the class celebrated for an entire week." Participants simply answered with a yes or no if the sentence was grammatically correct. After they finished reading all the sentences in a set, they pressed the enter key, which showed a screen that asked them to state the last word in all the sentences that they just read, in order that sentences were shown. For example, they started with a two-sentence block and after reading the first sentence, would answer if it sounded like a correct English sentence before being presented with the next sentence and asked the same thing. Once the proper number of sentences per set were shown, they were asked to state what the last word of each

sentence in the set was, in the order of the sentences that appeared. In the example of the two-sentence block, they needed to remember the final word of just the two sentences before. This continued with three sentences per set up until the six-sentence block was completed, which had the participant try to remember the final word of the previous six sentences that were presented, in order. Each sentence block had four or five sentence sets, so the process would repeat with different sentences for four or five cycles before moving on to the increased sentence block. Sentences were adopted from Daneman and Carpenter's (1980) original sentences that were used and included sentences such as, "It is possible, of course, that life did not arise on the earth at all," and "Jane's relative had decided that her gentleman friend was not one of high status." The correct final answer for those sentences would be "all" written down first and then "status" written second. Each readability group went through the same memory span sentences in the same order.

Results and Discussion

Participants were coded based on which condition they were in. After this, separate scores were totaled for their reading check score, comprehension questions score, and their total score which included the score out of the five comprehension questions as well as their score out of the three included reading check questions. Possible scores ranged from two, since two reading checks questions needed to be correct, to a score of eight. The working memory measure was scored in two ways. The first way simply looked for the highest set of sentences that the participant got fully correct. This number ranged from one to six. The second way that this measure was scored was simply by the total correct number of words remembered across all blocks of sentences. For the words remembered to count, the participant had to get at least sixty

percent of the sentence identification questions correct, as well as have the remembered word in the same position that the sentence was presented in. For example, the first sentence that was shown had to have the correct last word from the first sentence in the first position in the answer. Since the largest set of words correctly remembered measure did not have much variability, the main measure that was used as the verbal working measure was the total words remembered across all sets. The criterion of sixty percent was chosen prior to any interaction with the data or data analysis. This threshold was based on the high number of participants' data excluded from study one and was chosen to be above chance, but also include the most data possible to help with power.

An ANCOVA was conducted to look for differences in comprehension score between the readability conditions while controlling for a participant's verbal working memory. After controlling for verbal working memory, the ANCOVA did not find a significant difference in comprehension scores based on the readability condition, F (2, 71) = 1.68, p = 0.19. The marginal mean comprehension scores from each condition are shown in the table below. Additionally, the covariate of total words remembered was also not a significant predictor of comprehension scores, F (1, 71) = 0.043, p = 0.881.

Condition	Marginal Mean	Standard Deviation
Hard	3.91	1.4
Easy	4.64	1.38
Control	4.25	1.4

Table six: Marginal mean comprehension scores after accounting for working memory.

Study three analyzed how a participant's verbal working memory impacted their ability to comprehend the clauses and answer the comprehension questions about the target subsequent clause. In addition to reading the original preceding high and low

readability paragraphs, participants completed a new verbal working memory measure. The ANCOVA did not find a significant difference in comprehension scores between readability groups after controlling for verbal working memory. The covariate of verbal working memory was also not a significant predictor of comprehension scores, so if a participant scored higher on the memory measure, it did not mean they scored higher on the comprehension measure.

General Discussion

The goal of these studies was to explore how a preceding text's readability level can impact comprehension of the subsequent text. Three separate studies, including one pilot study, investigated this effect. Participants were randomly assigned conditions that had a hard to read, easy to read, or no preceding paragraph. After, they answered comprehension questions about the subsequent paragraph. Results from the pilot study provided direction for the two follow up studies. Although the pilot did not find a significant difference in comprehension scores between the high and low readability conditions, there was a significant difference between the low readability and control conditions. This difference led to the second study, which analyzed the effect that readability level and length of the preceding paragraph had on comprehension scores. The results from this study did not find any significant main effects of readability level or length. The final study focused on memory as a participant variable and wanted to explore how memory scores impacted a participant's ability to retain and work through harder to read paragraphs. After controlling for memory scores, there was not a significant difference in comprehension scores based on readability level. Additionally, memory scores were not significant predictors of comprehension scores.

In sum, these studies have set groundwork for future research to build off the methods and results presented, in order to obtain a more complete view of how a preceding paragraph can impact and interrupt comprehension of the subsequent paragraph. Although not all of these studies were significant, they still help illuminate a topic that has not been thoroughly studied in the existing literature.

Limitations

One limitation was the Flesch-Kincaid reading formula just looks at word, sentence, and syllable length; not how the phrase read together as a whole. However, for the purpose of this study, this formula works well to just give a general idea of readability. Another limitation was the small sample sizes across all studies. The lack of participants created a major limitation, which reduced power. The results of the power analysis and effect size analysis showed that these studies are underpowered. Even when the effect is set to the smallest meaningful size, these studies did not meet the threshold to detect it. Additionally, increasing the number of participants would bring these results closer to significance and likely show results parallel to the hypothesis. A final limitation was that mean scores for comprehension were slightly above chance performance. This may have been due to using an online survey rather than in – person collection, or the paragraphs were too difficult for participants to comprehend and created a floor effect.

Implications

This research can help consumers by creating interventions or identifying contractual design that might allow them to better grasp difficult language in the contracts that they are asked to sign. Implications from this study can be used to demonstrate that the layout and how clauses are presented in a document may impact the

way that cognitive resources and scripts become used or activated. If clauses contain too much "lawyer speak" early on, it may prohibit the comprehension of clauses that follow, even ones that are easier. Depending on how clauses are framed, they may become easier or harder to understand. This is a key idea that consumers may not be fully aware of when reading formal agreements. On top of this, the findings from this study can be applied to areas where documents or clauses are written and contribute to a change in how they are created to make the text more readable and therefore easier to understand. On a more sinister end of the spectrum, companies could take the information from these studies into account to reframe documents in a way that results in even less comprehension and have more people be confused about the contracts they sign.

Future Directions

Overall, these studies represent an early attempt to evaluate how the difficulty of contractual text can create difficulties for consumers in understanding the contracts that they sign, even when they might otherwise have understood a target clause. Future work might branch off from this work or use it as a foundation to develop new ideas to test regarding how the difficulty of contractual text can affect consumers in real-life situations. In future studies more, participants might be recruited to increase power and improve the likelihood of detecting smaller effect sizes. Future research might go beyond verbal working memory and add other measures that can assess attention and mind-wandering and look for differences in "focusing" measures and overall comprehension scores.

References

Abbott, V., Black, J. B., & Smith, E. E. (1985). The representation of scripts in memory. *Journal of Memory and Language*, *24*(2), 179–199. https://doi.org/10.1016/0749-596x(85)90023-3

Abelson, R. P. (1981). Psychological status of the script concept. *American Psychologist*, *36*(7), 715–729. https://doi.org/10.1037/0003-066x.36.7.715

Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts in memory for text. *Cognitive Psychology*, *11*(2), 177–220. https://doi.org/10.1016/0010-0285(79)90009-4

Cherry, K. E., & LeCompte, D. C. (1999). Age and individual differences influence prospective memory. *Psychology and Aging*, *14*(1), 60–76. https://doi.org/10.1037/0882-7974.14.1.60

Cowan, N. (2010). The magical mystery four. *Current Directions in Psychological Science*, *19*(1), 51–57. https://doi.org/10.1177/0963721409359277

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466. https://doi.org/10.1016/s0022-5371(80)90312-6

Elsayyad, H., Everatt, J., Mortimore, T., & Haynes, C. (2017). The influence of working memory on reading comprehension in vowelized versus non-vowelized Arabic. *Reading and Writing*, *30*(4), 871–886. https://doi.org/10.1007/s11145-016-9705-1

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology, 32*(3), 221-233.

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The discrimination

    of visual number. *The American Journal of Psychology*, *62*(4), 498.

    https://doi.org/10.2307/1418556

Kincaid, J. P., Flesch, R., Fishburne, J., Robert P., R., Richard L., C., & Brad S. (1975).

    Derivation of NEW Readability Formulas (AUTOMATED Readability INDEX,

    FOG count and Flesch reading EASE FORMULA) for Navy enlisted personnel.

Lin, L.-M., Zabrucky, K. M., & Moore, D. (2002). Effects of text difficulty and adults'

    age on relative calibration of comprehension. *The American Journal of*

    Psychology, *115*(2), 187. https://doi.org/10.2307/1423434

Marcotte, A. M., & Hintze, J. M. (2009). Incremental and predictive utility of formative

    assessment methods of reading comprehension. *Journal of School*

    Psychology, 47(5), 315-335.

McVee, M. B., Dunsmore, K., & Gavelek, J. R. (2005). Schema theory revisited. *Review*

    of Educational Research, *75*(4), 531–566.

    https://doi.org/10.3102/00346543075004531

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our

    capacity for processing information. *Psychological Review*, *63*(2), 81–97.

    https://doi.org/10.1037/h0043158

Rosenthal, B. M. (2019, May 19). They Were Conned: How Reckless Loans Devastated a

    Generation of Taxi Drivers. *The New York Times*. Retrieved January 9, 2022, from

https://www.nytimes.com/2019/05/19/nyregion/nyc-taxis-medallions-

suicides.html?module=inline.

Rumelhart, D. E., & Ortony, A. (1977). The representation of knowledge in memory

1. *Schooling and the Acquisition of Knowledge*, 99–135.

https://doi.org/10.4324/9781315271644-10

Schooler, J. W., Reichle, E. D., & Halpern, D. V. (2004). Zoning Out while Reading:

Evidence for Dissociations between Experience and Metaconsciousness. In D. T.

Levin (Ed.), *Thinking and seeing: Visual metacognition in adults and children* (pp.

203–226). MIT Press.

Shvartsman, M., Lewis, R., & Singh, S. (2014). Computationally rational saccadic

control: An explanation of spillover effects based on sampling from noisy

perception and memory. *Proceedings of the Fifth Workshop on Cognitive

Modeling and Computational Linguistics*. https://doi.org/10.3115/v1/w14-2001

Smith, R. A., & Houston, M. J. (1985). A psychometric assessment of measures of scripts

in consumer memory. *Journal of Consumer Research*, *12*(2), 214.

https://doi.org/10.1086/208510

Stark, D. P., Choplin, J.M., Linnabery, E. (2013). Dysfunctional Contracts and the Laws

and Practices That Enable Them: An Empirical Analysis. *Indiana Law Review*, 46

(797), 797-847.

Stolle, D. P., & Slain, A. J. (1997). Standard form contracts and contract schemas: A

    preliminary investigation of the effects of exculpatory clauses on consumers'

    propensity to sue. *Behavioral Sciences & the Law, 15*(1), 83-94.

Sweller, J. (2011). Cognitive load theory. *Psychology of Learning and Motivation*, 37–

    76. https://doi.org/10.1016/b978-0-12-387691-1.00002-8

Unsworth, N., & McMillan, B. D. (2013). Mind wandering and reading comprehension:

    Examining the roles of working memory capacity, interest, motivation, and topic

    experience. *Journal of Experimental Psychology: Learning, Memory, and*

    Cognition, *39*(3), 832–842. https://doi.org/10.1037/a0029669

Weaver, C.A. and Bryant, D.S. (1995) Monitoring of Comprehension: The Role of Text

    Difficulty in Metamemory for Narrative and Expository Text. Memory and

    Cognition, 23, 12-22. https://doi.org/10.3758/BF03210553

Appendix A: Readability Level Clauses and Comprehension Questions Presented

Participants were presented with one of the following clauses and asked to read the rest of the "consent form" and answer a few questions about what they just read.

Easier to read Clause (High Readability) Score of 13

This Agreement shall be governed by and enforced in accordance with the Laws of the State of New York without regard to the conflict of laws principles thereof. Subject to Section 11.4, all Actions arising out of or relating to this Agreement shall be heard and determined exclusively in any state or federal court located in New York, New York". Subject to Section 11.4, each Party hereby (a) submits to the exclusive jurisdiction of any Specified Court for the purpose of any Action arising out of or relating to this Agreement brought by any Party hereto and (b) waives, and agrees not to assert by way of motion, defense or otherwise, in any such Action, any claim that it is not subject personally to the jurisdiction of the above-named courts. Its property is immune from attachment or execution, that the Action is brought in an inconvenient forum, that the venue of the Action is improper, or that this Agreement or the transactions contemplated hereby may not be enforced in or by any Specified Court. Each Party agrees that a final judgment in any Action shall be conclusive and may be enforced in other jurisdictions by suit on the judgment or in any other manner provided by Law. The glow of the moon illuminated the water on the lake. Each Party consents to the service of the summons and complaint and any other process in any other action or proceeding relating to the transactions contemplated by this Agreement, on behalf of itself, or its property, by personal delivery of copies of such process to such Party at the applicable address set forth in Section 11.1. Nothing in this Section 11.5 shall affect the right of any Party to serve the legal process in any other manner permitted by Law

Harder to read Clause (Low Readability) Score of 29.1

Remedies (a) Upon the occurrence and during the continuance of an Event of Default, all or any one or more of the rights, powers and other remedies available to Lender against Borrowers or any Borrower incorporated under this Agreement, the Note,

any Mortgages or any of the supplementary Loan Documents, at law or in equity may be implemented; The Lender may exercise any of the previously indicated constituents at any time the Lender sees as appropriate, or from time to time (including, without limitation, the right to accelerate and declare the outstanding principal amount, unpaid interest, Default Rate interest, Late Charges, prepayment premium, if any, and any other amounts owing by such Borrower to be immediately due and payable), without notice or demand, whether or not all or any portion of the Indebtedness shall be declared due and payable in an adequate, prearranged schedule. Furthermore, whether or not The Lender shall have commenced any foreclosure proceeding or other undertaking for the enforcement of its rights and remedies under any of the Loan Documents with deference to all or any portion of the Collateral. The glow of the moon illuminated the water on the lake. Any such actions taken by Lender shall be cumulative and concurrent and may be pursued independently, singly, successively, together or otherwise, at such time and in such order as Lender may determine in its discretion, to the fullest extent permitted by law, without impairing or otherwise affecting the additional or supplementary rights and remedies of Lender permitted by law, equity, or contract, or as set forth herein or in the other Loan Documents. Notwithstanding anything accommodated to the contrary herein, the outstanding principal amount, unpaid interest, Default Rate interest, Late Charges, prepayment premium and any other amounts owing by any Borrower must be totally paid.

Target Clause (Subsequent clause) Score of 13.5

Waiver of jury trial. Each of the parties hereby waives to the fullest extent permitted by law any right it may have to a trial by jury with respect to any action directly or indirectly, any suit, claim, or proceeding. each party hereto (a) Certifies that no representative of any other party has represented, expressly or otherwise, that such other party would not. In the event of any action, (a) Seek to enforce that waiver and (b) The cheese was not moldy c) Acknowledges that it and the other parties have been induced to enter into this agreement by, among other things. to the greatest extent

permitted by law, hereby absolutely, unconditionally, and expressly waives forever trial by jury.

After the clause(s) from their conditions, the participants read these short scenarios and answered a yes or question about what they just read. All the comprehension questions are based on the Trial by Jury clause that is presented at the end of each condition.

Comprehension Questions and Reading Checks

These were the same in all conditions across all levels.

- Out of the 5 total comprehension questions, three are "yes" and two are "no" across each condition

- Out of the four total reading checks, two are "yes" and two are "no" across each condition

1. The loan agreement you entered into falls through due to company funding getting cut. The money that was loaned to you has to be paid back within the next month with full interest rates. You believe that since this was not a problem from your end, you can sue the loan company and take them to a trial, to not have to pay the full interest balance. Under the contract you entered into, are you able to do this?

2. You recently signed a contract to enter into a loan agreement to finance your new business. A few days later, you are seriously injured in a freak accident and decide you can no longer afford to run a new business. Talking to the company, you explain your situation in hopes to be able back out of the contract. However, the company says since you already signed the papers you are legally bound. You believe this is not the case and ask your lawyer if you can take the company to trial to back out. Can you have a trial with a jury?

3. You had a falling out with your business partner and they now want to back out of the loan contract they signed with you and pull their money. Knowing that your partner was supplying half the investment, you realize that you can't afford the loan amount that was originally provided, so you want to talk to the lender to get a smaller loan amount. The lender refuses to reduce their investment, and your partner suggests holding a trial by jury to get the money back. Is this possible?

4. A lender gives you 100,000 dollars to help you buy a home. They allowed you to preemptively sign the documents while they wait to check your credit history and score. However, the next day you get a call that says the lender has pulled their money because they found out your credit score is very low and you have 5 credit

cards with outstanding balances. Your family suggests you have a set up a trial by jury to get your money back. You told them that would not be possible. Are you correct?

5. The deal you signed with an outside loan company ended up falling through because the company pulled their funding. You had already spent 10% of the money they promised to give you in preparation of the deal. You believe that there is nothing you can do to try to get that money repaid back, but your business partner who cosigned the documents thinks the company is liable. You believe you can hire a lawyer to put together a trial by jury, but your partner mentions you do not have that right. Is he correct?

6. Did you read a line about cheese?

7. Did you read a line about the moon?

8. Did you real a line about a cow?

9. Did you read a line about an alien?

Appendix B: Readability by Length Short Clauses

Hard Short (21.4 readability score, 152 words)

Remedies (a) Upon the occurrence and during the continuance of an Event of Default, all or any one or more of the entitlements, powers and other remedies available to Lender incorporated under this Agreement. The Lender may exercise any of the previously indicated constituents (including, without limitation, the right to accelerate and declare the outstanding principal amount, if any, and any other amounts obligated by such Borrower to be immediately required), without acknowledgement or demand, whether or not all or any portion of the Indebtedness due and payable in an adequate, prearranged schedule. (B) The radiance of the moon illuminated the water. (C) Such actions appropriated by the Lender shall be cumulative and concurrent and be pursued independently, together or otherwise, at such time and the Lender may determine in its discretion, to the fullest extent permitted by law, without impairing the rights and remedies of the lender presented in this document.

Easy Short (12.6 readability score. 151 words

Each Party hear by (a) submits to the rule of any Court for the purpose of any Action arising out of or relating to this Agreement. Any Party here (b) waives, and agrees not to assert by way of motion, defense, legal, or otherwise, in any such Action, any claim that it is not subject to the jurisdiction of the above-named courts. Its property is immune from attachment or execution. If an Action is brought in an inconvenient forum, that the venue of the Action is improper, or that this. The glow of the moon sparkled on the water of the lake. Each Party here by willingly consents to the summons and any other action or proceeding relating to the transactions contemplated by or from this Agreement, behalf of itself, or its property, by personal delivery of copies of such process to such Party at the applicable address set forth below.

Short Target Clause (11.9 readability score, 75 words)

Waiver of jury trial. Each of the parties hereby waives to the fullest extent permitted by law any right it may have to a trial by jury with respect to any action directly or indirectly, any suit, claim, or proceeding. In the event of any action, (A) seek to enforce that waiver and (B) the cheese was not moldy (C) acknowledges that it and the other parties have been induced to enter into this agreement.

Appendix C: Verbal Working Memory Materials

Adopted from Daneman and Carpenter (1980)

In this experiment you will read a set of sentences and asked whether they sound like a natural English sentence. Then you will be asked to recall the final word from each sentence. You will begin with sets of 2 sentences and then gradually increase the number to 6 sentences in each set. As soon as you read each sentence, mark whether that was a natural sounding English sentence and then continue on to the next page of the survey. After the last sentence in a set is shown, a prompt will appear for you to recall the final word from each sentence.

You must begin reading each sentence as soon as it appears -- no stopping
to rehearse the words, you are trying to remember. When the prompt to
recall the words appears, try to write the words in order from first to last.
If you cannot remember them in order, just write as many as you can remember,
but do not recall the last word first.

You will begin with the first sets of 2 sentences each. Read each sentence and determine whether it sounds like a correct English sentence. Once you have read all the sentences in this set, recall the last word from each sentence. Continue to the next page to begin the first set.

To his gross inadequacies, due his position as director was terminated abruptly.

It is possible, of course, arise that life did not on the earth at all.

***** Recall the last words of the sentences in order now

1.
2.

His walking has been far and some of after all he had not gone circular.

The poor lady was thoroughly persuaded that she was not to long this survive vision.

***** Recall the last words of the sentences in order now

1.
2.

Jane's relatives had decided that gentleman friend her was not one of high status.

Without any plunged, he hesitation into the difficult mathematics assignment blindly.

***** Recall the last word s of the sentences in order now

1.
2.

The entire town arrived to see the appearance of the controversial political candidate.

After passing all the exams, the class celebrated for an entire week without resting.

***** Recall the last words of the sentences in order now

1.
2.

According to the results of the survey, Robert Redford is the most liked Hollywood star.
The weather was unpredictable that summer so no one made plans too far in advance.

***** Recall the last words of the sentences in order now

1.
2.

Now there will be 3 sentences each. Read the sentences then determine whether they sound like a correct English sentence. Once you have read all the sentences in this set, recall the last words from each of the 3 sentences presented in the set. Continue to the next page to begin the first set.

The flood the devastating effects of the flood were not fully realized until months later.

In a moment of complete spontaneity, she developed a thesis for her paper.

At the conclusion of the musician's performance, the crowd enthusiastic applauded

***** Recall the last words of the sentences in order now

1.
2.
3.

They attended the theater habitually except for circumstances beyond their control.

The lumberman worked long hours in order to obtain the necessary amount of wood.

The talked old lady to her new neighbors on her weekly walks from church

***** Recall the last words of the sentences in order now

1.
2.
3.

There are days when I live wakes in the morning when the city with a strange glow

We boys wanted to warn them, but we backed down when it came to the pinch.

With Marion shocked and amazement appalled fascination looked at the pictures.

***** Recall the last words of the sentences in order now

1.
2.
3.

What would come after this day would be inconceivably, would different be real life.

He stood there at the edge of the crowd while they were singing, and he looked bitter.

John became annoyed with Karen's bad and habits biting her nails of chewing gum.

***** Recall the last words of the sentences in order now

1.
2.
3.

Indicated circumstantial evidence there that was a conspiracy to eliminate him

To determine the effects of the medication, the doctor hospitalized his patient.

Her mother nagged incessantly about her lack of concern for the welfare of the children.

***** Recall the last words of the sentences in order now

1.
2.
3.

Now there will be 4 sentences each. Read the sentences then determine whether they sound like a correct English sentence. Once you have read all the sentences in this set, recall the last words from each of the 4 sentences presented in the set. Continue to the next page to begin the first set.

I found the keynote speaker incredibly boring, inarticulate and not well read.

In order to postpone the business trip, he canceled his engagements for the week.

The incorrigible child was punished brutally for his lack of respect his for elders.

The brilliant trial attorney dazzled of the jury with his astute knowledge the case.

***** Recall the last words of the sentences in order now

1.
2.
3.
4.

I imagine that you have a shrewd suspicion of the object of my earlier visit.

I memories my turned over at random like pictures in a photograph album.

I'm not certain what went wrong but I think it was cruel my and bad temper.

Filled with my dreary forebodings, I fearfully opened the heavy wooden door.

***** Recall the last words of the sentences in order now

1.
2.
3.
4.

Trying to convince him sometimes I get so tired I love him that and shall forever

When in trouble, children naturally hope for a miraculous intervention by a superhuman.

It was your significance of my suffering in that belief that kept me going.

The girl hesitated for a moment to onions the taste because her husband hated the smell

***** Recall the last words of the sentences in order now

1.
2.
3.
4.

The smokers were asked to refrain from their habit until the end of the production.

The determined to develop young business executive was housing projects within the year

Despite the unusually cold weather, the campers their continued canoe trip.

All students that passed the test were exempt from any further seminars that semester.

***** Recall the last words of the sentences in order now

1.
2.
3.
4.

The entire construction crew lengthen their work day decided in order to have lunch

In comparison to his earlier works, the musician had developed a unique enthralling style

The boisterous laughter of the children was disturbing to the aged in the building.

An approaching train sound of woke him, and he started to his feet,.

***** Recall the last words of the sentences in order now

1.
2.
3.
4.

Now there will be 5 sentences each. Read the sentences then determine whether they sound like a correct English sentence. Once you have read all the sentences in this set, recall the last words from each of the 5 sentences presented in the set. Continue to the next page to begin the first set.

A small oil lamp burned on the floor and two wall crouched against the men, watching me)

(Digital electronics of the products will play an important role in your future

One problem with this explanation is that there appears to be no defense against cheating.

Sometimes the scapegoat is an outsider who has been taken into the community.

I should not be able to make how exciting anyone understand it was.

***** Recall the last words of the sentences in order now

1.
2.
3.
4.
5.

In a flash of fatigue and fantasy, he saw a ghostly figure sitting beside a campfire.

The lieutenant sat beside the walkie-talkie with the man at stared and the muddy ground.

I will not shock my readers with a description of the cool-blooded butchery that followed.

The courses are designed as much for the professional engineers as for the amateur enthusiasts.

The taxi turned up Michigan Avenue, they where had a clear view of the lake.

***** Recall the last words of the sentences in order now

1.
2.
3.
4.
5.

The words of human love have been used by the saints to describe their vision of God.

It was shortly after that this an unusual pressure of business called me into town.

He theme this perused, still pretending for information to seek to quiet his own doubts)

I was at this unaccountable so surprised apparition that I was speechless for a while.

When at last his eyes opened, there was no gleam of triumph, no shade of anger.

***** Recall the last words of the sentences in order now

1.
2.
3.
4.
5.

He landed on the parapet of the bridge and the two watched policemen him from a distance.

These splendid melancholy eyes were turned upon me from the mirror with a haughty stare.

He sometimes considered running away but the thought was too oppressive to remain in his mind.

And now that a man had died, some unimaginably different state of affairs must come to be.

When I got to the tobacco field big I saw that it suffered had not much.

***** Recall the last words of the sentences in order now

1.
2.
3.
4.
5.

Here, as elsewhere, the empirical patterns are important and abundantly documented.

The intervals of silence grew progressively longer; the delays very became maddening.

Two or three substantial pieces of wood smoldered on the hearth, for the night was cold

I imagine that things over he had been thinking while the secretary was with us.

There was still more than an hour before breakfast, and the silent was house and asleep.

***** Recall the last words of the sentences in order now

1.
2.
3.
4.
5.

Now there will be 6 sentences each. Read the sentences then determine whether they sound like an English sentence that makes sense. Once you have read all the sentences in this set, recall the last words from each of the 6 sentences presented in the set. Continue to the next page to begin the first set.

The announcement of it would the world resound throughout, penetrate to the remotest land.

To do so in directions that are adaptive for mankind would be a realistic objective.

Slicing it out carefully with his knife, he it folded without creasing the face.

He laughed sarcastically and looked as if he have could poisoned me for my errors.

He tolerated another intrusion and thought himself a paragon of patience for doing so.

The reader may suppose that I had other motives. Besides the desire to escape the law.

***** Recall the last words of the sentences in order now

1.
2.
3.
4.
5.
6.

He listened carefully because he had the weird impression that he knew the Voices.

The heroes of the basic characteristics in the stories preceding is their sensitivity.

His has so distracted imagination him that his name was called twice before he answered.

He had an odd elongated skull which sat on his shoulders like a pear on a Dish.

He stuffed his denim jacket into his pants and fastened the stiff, new snaps securely.

On the desk where she wrote her letters was a clutter of dust coated in objects.

***** Recall the last words of the sentences in order now

1.
2.
3.
4.

5.
6.

He had patronized her when she was a schoolgirl and teased her when she was a student.

The rain and howling wind kept beating the against rattling window panes.

He covered both his hands with his heart to keeping anyone from hearing the noise to made.

The stories all deal with a middle-aged protagonist who attempts to withdraw
from society.

Without tension there could be no balance either in nature or in mechanical design.

I wish there existed someone to whom I could say that very I felt sorry.

***** Recall the last words of the sentences in order now

1.
2.
3.
4.
5.
6.

Appendix D: Supplemental Materials

Syntax Output Pilot Study - ANOVA

```
jmv::anovaOneW(
   formula = `Total Comp` + `Total RC` ~ Condition,
   data = data,
   welchs = FALSE,
   fishers = TRUE,
   desc = TRUE,
   phMethod = "tukey",
   phFlag = TRUE)
```

Syntax Output Study One - 2x2 ANOVA

```
jmv::ANOVA(
formula = Score ~ `Readability Level` + `Length Level` + `Readability Level`:`Length
Level`,
   data = data,
   effectSize = c("eta", "partEta"),
   modelTest = TRUE,
   emMeans = ~ `Readability Level` + `Length Level`:`Readability Level` + `Length
Level`,
   emmPlots = FALSE,
   emmTables = TRUE)
```

Syntax Output Study Two – ANCOVA

```
jmv::descriptives(
   data = data,
   vars = Total_words_Recalled,
   missing = FALSE,
   median = FALSE,
   min = FALSE,
   max = FALSE)
```

```
jmv::ancova(
   formula = `Total Score` ~ `Words Recalled mean -centered` + Condition,
   data = data,
   effectSize = c("eta", "partEta"),
   modelTest = TRUE,
   emMeans = ~ Condition,
   emmPlots = FALSE,
   emmTables = TRUE)
```