Voices from Academia

Minimizing the Complexity of Public Health Documents: Making COVID-19 Documents Accessible to Individuals Who Read Below the Third-Grade Level

Sofia Benson-Goldberg, Ph.D.¹, Lori Geist, Ph.D.¹, Ben Satterfield, Ed.D.², Deidra Bunn, SLP², and Karen Erickson, Ph.D.¹

¹Center for Literacy and Disability Studies, Department of Allied Health Sciences
University of North Carolina at Chapel Hill
²Center for Inclusive Design and Innovation
Georgia Institute of Technology

Corresponding Author
Sofia Benson-Goldberg
University of North Carolina-CH
CB #7335 UNCH-CH
Chapel Hill, NC 27599-7335
Email: sofia_bensongoldberg@med.unc.edu

ABSTRACT

The Centers for Disease Control and Prevention (CDC) is a trusted source for public health information, but people must be able to access and understand that information for it to be used. The CDC and the CDC Foundation recognized the need to ensure that its guidance documents related to COVID-19 were accessible to the full range of individuals with disabilities, including those with intellectual and developmental disabilities who read or listen with comprehension at or below the third-grade level. In response to this need, they contracted with the Center for Literacy and Disability Studies (CLDS), Department of Allied Health Sciences, University of North Carolina at Chapel Hill, and the Center for Inclusive Design and Innovation, Georgia Institute of Technology, to create easy to read versions of a collection of guidance documents related to COVID-19. The CLDS began the process by seeking existing guidelines or research to support the creation of these documents. When no such information was
located, the CLDS conducted a systematic review of the literature and developed the *Minimized Text Complexity Guidelines*. The outcomes and benefit of this work include improved access to critical information regarding COVID-19 for individuals with intellectual and developmental disabilities, as well as other adults who read and listen with comprehension below a third-grade level.

**Keywords:** accessible text, health literacy, low-literacy resources

**MINIMIZING THE COMPLEXITY OF PUBLIC HEALTH DOCUMENTS: MAKING COVID-19 DOCUMENTS ACCESSIBLE TO INDIVIDUALS WHO READ BELOW THE THIRD-GRADE LEVEL**

An estimated 56 million adults in the United States have basic or below-basic reading skills (Lesgold & Welch-Ross, 2012). While these adults are able to read simple words and phrases in familiar contexts, they are largely unable to do so with comprehension. Simultaneously, more than 61 million people in the United States live with a disability. Reading comprehension is consistently the academic area with the lowest outcome for all adults with disabilities (Wagner et al., 2006). Within this group, adults with intellectual disabilities are known to struggle with reading in general and reading comprehension in particular, with the population consistently scoring nearly three standard deviations below the mean on reading comprehension measures (Wagner et al., 2006).

Though poor reading comprehension has long been associated with negative outcomes for adults with and without disabilities (Johnson, 1985; UNESCO, 2005), it posed a significant threat during the COVID-19 pandemic. Most information regarding COVID-19 safety precautions was produced and disseminated in print formats. Unfortunately, the complexity of many of these materials was beyond the reading comprehension level of many people with intellectual and developmental disabilities. This presented a significant problem during the COVID-19 pandemic, as people with intellectual disabilities who contracted the virus were three (Rabin, 2020) to eight (Gleason et al., 2021) times more likely to die than those without intellectual disabilities. The risks of getting sick and dying from COVID-19 are compounded for people with intellectual disabilities from racial and ethnic minority groups who face additional inequities in social determinants of health (CDC, 2020). Lack of access to COVID-19 information due to difficulty comprehending text likely contributed to this. In general, limited access to health information has negative health implications and shines a light on the underlying factors contributing to the increased risks faced by people with intellectual disabilities in the current pandemic.

Section 508 of the Rehabilitation Act mandates that public-facing electronic communication content created by any Federal agency must be accessible for individuals with some disabilities (e.g., assistive technology devices can access information, closed captions are available on videos, etc.). However, the requirements fall short of ensuring that such communication content is accessible to all individuals with disabilities. One specific group that has been overlooked is people with and without intellectual and developmental disabilities who have extremely low literacy (i.e., reading and listening comprehension at the third-grade grade level or lower). Though plain language standards do exist (see...
https://www.plainlanguage.gov/guidelines), there are currently no concise guides for instructing federal agencies on how to expand communication reach and accessibility to include those people who read with comprehension below a third-grade level.

In response to this problem, and in order to create a set of easy to read documents based on the COVID-19 guidance documents produced and disseminated by the Centers for Disease Control and Prevention (CDC), a team from the Center for Literacy and Disability Studies (CLDS) Department of Allied Health Sciences at the University of North Carolina at Chapel Hill created the Minimized Text Complexity (MTC) Guidelines (Erickson et al., 2020) to inform writing text that is understandable to people with disabilities who read at or below a third-grade level. After developing the guidelines, the team at the CLDS worked in collaboration with a team from the Center for Inclusive Design and Innovation, Georgia Institute of Technology, to apply them to author a set of simplified, easy to read documents related to COVID-19.

TARGET AUDIENCE

There are two primary audiences for the work described here. The first audience includes the 56 million adults in the United States who have below-basic reading levels (Lesgold & Welch-Ross, 2012), including the majority of adults with moderate to severe intellectual and developmental disabilities. Our goal in creating the MTC Guidelines and applying them to the CDC’s guidance regarding COVID-19 was to make the documents accessible to this entire group of individuals. Our secondary audience includes everyone who has a role in creating health and other documents for people in our primary audience. This secondary audience includes a broad coalition of people including: (a) people who devote their lives to communicating critical health information to the public; (b) administrators and other personnel in healthcare and education settings who occasionally prepare documents to communicate critical information; (c) researchers who are required to make their consent documents and other materials accessible to the people who participate in their research; and (d) anyone who wants to create cohesive and easy to read information for adults who read or listen with comprehension at a third-grade level or lower.

LITERATURE REVIEW

Plain Language Due

Since the passing of the Plain Writing Act of 2010, the law now requires that all federal agencies use “clear Government communication that the public can understand and use” (p. 1). There are official plain writing guidelines that accompany the Plain Writing Act (https://www.plainlanguage.gov/guidelines). The guidelines are broadly organized around audience considerations, organizational practices, and writing principles; with writing principles further broken down to word, sentence, paragraph, and section-level guidelines. Writers are encouraged to use lists, tables, bullets, illustrations, and headings to make complex materials easier to understand. Word-level guidelines include using (a) active voice (b) contractions, (c) pronouns, and (d) short, simple words.
Explicit training is necessary to support experts in using plain language in their written and oral communications; furthermore, even with training, experts require multiple passes back and forth with another expert in order to achieve sufficiently easy texts (Hadden, 2015). This suggests that much time and attention is required in order to achieve texts that are compliant and consistent with plain language guidelines. Widespread adoption has been found when easily accessible templates and procedures are provided for writers unfamiliar with how to implement plain language guidelines (Hadden et al., 2017). While plain language guidelines generally result in texts that have readability scores between sixth and the eighth-grade levels, in certain contexts, plain language may be insufficient to lead to comprehension (Miles & Cottle, 2011).

**Text Complexity**

Text complexity refers to distinct properties of a text that exist regardless of reader or task (Mesmer et al., 2012). Properties that impact complexity include vocabulary, word density, sentence length, and text length (Hiebert, 2014). Importantly, this is different from text difficulty, which refers to how easy or difficult a text is for a given reader. Text difficulty differs from person to person, and task to task. In fact, difficulty might be viewed as a result of the relationship between the person, the task, and the text (McNamara et al., 1996).

Text complexity impacts learning from text in a number of ways. When text complexity is lower, people demonstrate greater comprehension of material (Treptow et al., 2007). For example, children have been shown to learn more new vocabulary words when reading texts that are less complex (Anderson, 1996). Furthermore, regardless of ability, people learn and retain more new information from expository texts when the text complexity level is lower than the level of text they typically read with comprehension (Allington et al., 2015). Given the demands to learn novel words and to retain new information to navigate COVID-19 guidance, minimizing text complexity was of the utmost importance.

Readability formulas such as Flesch-Kincaid and Lexile scores are often used to identify texts that might be easier to read. While these systems are helpful for sorting large groups of text, they provide little information about what specifically is impacting complexity in any given text (Hiebert, 2011; Hiebert, 2014). Furthermore, these formulas privilege text features that often fail to reduce complexity. For example, sentence length weighs heavily in these formulas, with texts with shorter sentences receiving lower readability scores than texts with longer sentences. This is problematic because shortening sentences does not necessarily improve comprehension (Arya et al., 2011), and in fact can impair comprehension (Beck et al., 1982), presumably because it interferes with cohesion (Graesser et al., 2004). In order to establish guidelines for producing texts that can consistently be read with comprehension by the target population, it was important to go beyond readability formulas, and identify specific text features that impact complexity.

**Graphics, Icons, and Images**

Pairing graphics with text is often assumed to be a way of creating resources that demand fewer cognitive skills and thereby help readers extract meaning from text, presumably by eliminating the need to decode words. As such, the recommendation is often made to add pictures, pictograms, or graphic symbols to
text in order to make it easier to read (e.g., Nomura et al., 2010; Office for Disability Issues, 2018). Adding pictures, pictograms, or graphic symbols might involve representing the meaning of an entire sentence, key ideas in a sentence, or all of the meaning and grammatical information in a sentence (Poncelas & Murphy, 2007). Unfortunately, there is little empirical support for any of these uses of pictures, pictograms, or graphic symbols to support text comprehension (Benson-Goldberg & Erickson, 2020; Erickson et al., 2010; Hurtado et al., 2014).

Pairing graphic symbols with individual words has been used successfully in the context of expressive communication for many decades (e.g., Samuels, 1967); however, use in the context of learning to read (Erickson et al., 2010) and supporting comprehension (e.g., Hurtado et al., 2014; Poncelas & Murphy, 2007) is limited. Nonetheless, guidelines regarding text accessibility continue to include recommendations for the inclusion of graphic symbols (e.g., Department of Health, 2010; Office for Disability Issues, 2018), and various stakeholders respond positively and enthusiastically to texts that are supported word-by-word with graphic symbols (Parson & Sherwood, 2015).

Other visual supports, such as infographics, pictograms, and pictures, are often assumed to support reading comprehension. While it is true that readers rate infographics as more user-friendly, comprehension assessments do not demonstrate an advantage for infographics (Buljan et al., 2018). Furthermore, the research suggests that adding any visual support to text makes reading with comprehension more difficult, as it not only requires the reader to alternate their visual attention between the text and the other visual information, but it also requires different neural mechanisms to process and understand the information (Fisher & Frey, 2014; Hegerty et al., 1991). When pictures or illustrations have been used successfully, they have been closely matched to explanatory text that was clearly understood by the reader (Houts et al., 2006). Unfortunately, beginning and very low-level readers who do not easily understand the text pay attention to unrelated details in images or otherwise attend to the pictures in ways that do not support their comprehension (Filippatou & Pumfrey, 2006).

Given the need to concentrate on maximizing comprehension of the COVID-19 guidance documents, it was important to consider the ways in which adding pictures, graphic symbols, and infographics to text often increase rather than decrease the text complexity, and interfere with comprehension. Careful attention was paid to the literature regarding use and placement of pictures, graphic symbols, and infographics in order to determine how best to leverage these in order to support comprehension and minimize complexity.

**METHODS**

Rapid reviews are necessary to strengthen health policy and systems during time-sensitive events such as the COVID-19 global pandemic presents (Tricco et al., 2017). Similar to traditional systematic reviews, rapid reviews involve comprehensive reviews of the research literature with rigorous search and review strategies. Different from traditional systematic reviews, the goal is to progress rapidly to actionable understandings. The rapid review that informed the development of the MTC Guidelines involved a search of electronic databases including PsychInfo, MEDLINE, ERIC, CINAHL, Health Source:
Nursing/Academic Edition, and Education Full Text for records published within the last 10 years. These databases were chosen to get results from various subject areas. For studies to be included in the resulting guidance, they needed to address ways in which text could be written to support comprehension for any beginning or struggling readers, including those with developmental or intellectual disabilities.

**Search Strategy**

Initial searches were conducted in July 2020. Databases were searched using the keywords “plain language” or “simplified text” or “easy to read.” The term “text complexity” was then added because it is closely tied to research on reducing text difficulty at the lowest text levels. The use of these search terms resulted in 1,716 records that were uploaded to Covidence. After duplicates were removed, 1,304 records remained. Each was screened for relevance to the problem at the title and abstract level and 92 records were retained that related to increasing comprehension by using plain language or common, everyday words and excluded records on unrelated topics ($n = 1,212$). The 92 relevant records were then reviewed at the full-text level for appropriate recommendations. The search process is illustrated in Figure 1.

![Figure 1: Summary of Review Results](image)

**RAPID REVIEW RESULTS: INITIAL GUIDELINES**

The MTC Guidelines were written by examining the 92 retained records, as well as recent work completed regarding text complexity and beginning readers that was conducted by Cunningham et al. (2005) and Schuster and Erickson (2014). Records were read and examined to determine what text features are essential to supporting text comprehension for beginning and struggling readers.

The review of the resulting records indicated that plain language ([https://plainlanguage.gov](https://plainlanguage.gov)) was a necessary precondition of simplifying text, but insufficient to ensure comprehension for people with text comprehension skills below the sixth-grade level. Furthermore, it became clear that simply relying on
readability metrics would be insufficient to result in texts that would simultaneously support comprehension and retain the meaning that is critical to the COVID-19 guidance (Cunningham et al., 2018; Krieger et al., 2017; Leroy et al., 2013). While readability formulas are helpful for estimating the difficulty of texts relative to each other, they are insufficiently sensitive to syntactic and semantic features of text, as well as format features that have significant impacts on reading comprehension (DuBay, 2004; Krieger et al., 2017; Leroy et al., 2013).

As a result, it was important for the MTC Guidelines to focus on whole-text level guidelines, as well as specific sentence- and word-level guidelines that would support writers in producing texts that could be read easily, while still supporting comprehension of new information. The MTC Guidelines were organized into four sections: (a) whole-text-level guidelines; (b) sentence-level guidelines; (c) word-level guidelines; and (d) document-format guidelines. See https://www.med.unc.edu/ahs/clds/resources for full guidelines.

When compared to plain language guidelines, the MTC Guidelines differ in several important ways. First, the word-level guidelines are more specific and stringent. Where plain language suggests that authors should choose “short words,” the MTC Guidelines suggest that authors should ensure that at least 92% of words are amongst the most frequently occurring in written English. Furthermore, the MTC Guidelines require authors to reduce syllables per word, rather than word length in general. Similarly, content guidelines are more stringent in the MTC Guidelines. Where plain language guidelines begin by encouraging the writer to “think about the audience,” the MTC Guidelines go further by requiring that the document presents information that is directly relevant to the lives of the intended audience. If the content is not directly relevant, it should be excluded. Second, there are several plain language guidelines that were specifically contradicted in the MTC Guidelines. For example, plain language guidelines encourage using contractions, negation, and many useful headings, including questions, statements, and topics. None of these recommendations appear in the MTC Guidelines. In fact, contractions and negations are to be avoided entirely. Furthermore, headings are to be single informative statements of eight words or less. Third, where plain language guidelines make the general suggestion that authors design documents for easy reading, MTC Guidelines offer specific recommendations in regard to layout and organization for the target population.

APPLICATION OF GUIDELINES TO CDC COVID-19 EMERGENCY RESPONSE MATERIALS

The MTC Guidelines were applied to a total of 25 documents, prioritized by project staff from the CDC Foundation (CDC-F) and the CDC. The goal was to rapidly produce documents that conveyed the content deemed most critical within each document to help the target audience safely navigate the COVID-19 pandemic. Documents were extracted directly from html links provided by the CDC. The text was copied into worksheets, with text sectioned by content. This text was then annotated to indicate which MTC Guidelines needed to be applied to every section of text. At least two researchers trained in the application of MTC Guidelines then authored simplified versions of each section of text on the worksheet.
These solutions were compiled and compared by the first and last authors in order to arrive at documents that conveyed the most critical content, while adhering to the MTC Guidelines.

Results from First Ten Documents

The first ten documents simplified using the MTC Guidelines were reviewed for accuracy and relevance by project staff at the CDC-F and then cleared by appropriate offices at the CDC before they were posted on the CDC website. The ten documents can be accessed from the CDC website (https://www.cdc.gov/coronavirus/2019-ncov/easy-to-read/index.html) and a microsite created by the Center for Inclusive Design and Innovation at Georgia Tech (https://cidi.gatech.edu/covid/easytoread). The ten documents included in the analysis that follows are listed in Table 1.

<table>
<thead>
<tr>
<th>Document Title</th>
<th>URL</th>
</tr>
</thead>
</table>

The following metrics were tabulated for both the original and the documents simplified using the MTC Guidelines: (a) total word count; (b) words per sentence; (c) syllables per word; (d) type token ratio; (e) percent of words appearing among the 3,000 most frequently occurring words in written English; (f) reading ease; and (g) grade level. Percent of words appearing among the 3,000 most frequently occurring words in written English was determined using search tools at https://www.wordandphrase.info/analyzeText.asp. The remaining metrics were determined using SEO Scout tools at https://seoscout.com/tools/text-analyzer. Use of both sets of online tools required that the text be formatted in particular ways to ensure accuracy and consistency. Specifically, hyphens were deleted so that COVID-19 would be treated as a single word. In addition, text had to be entered as a
single paragraph, with all text in lowercase and all punctuation removed to determine the percent of words appearing among the 3,000 most frequently used words in written English.

When taken together, the first ten source documents from the CDC had 12,654 words with an average readability of 8.4 (+ 2.3) grade level as determined by the SEO Scout site using a combination of the Flesch Kincaid reading ease algorithm, supplemented with the SMOG index, Automated Readability Index, Gunning Fog, and Coleman Liau scores (see https://seoscout.com/tools/text-analyzer for more). After being simplified following the MTC Guidelines to convey only the information deemed pertinent to the target population, the ten simplified documents had 4,657 words with an average readability of 2.4 (+ 0.5) grade level. Table 2 displays the average metrics for the first ten documents in their original and simplified forms.

<table>
<thead>
<tr>
<th>Table 2: Average Metrics for the Ten Original and Simplified Documents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metric</td>
</tr>
<tr>
<td>Total Words( ^1 )</td>
</tr>
<tr>
<td>Average Number of Words per Sentence( ^1 )</td>
</tr>
<tr>
<td>Average Number of Syllables per Word( ^1 )</td>
</tr>
<tr>
<td>Type Token Ratio (Unique Words/Total Words( ^1 )</td>
</tr>
<tr>
<td>Percent of Words among 3,000 Most Frequent( ^2 )</td>
</tr>
<tr>
<td>Flesh Kincaid Reading Ease( ^1 )</td>
</tr>
<tr>
<td>Grade Level Equivalent on Readability Index( ^1 )</td>
</tr>
</tbody>
</table>

\( ^1 \)Metric generated at https://seoscout.com/tools/text-analyzer

\( ^2 \)Metric generated at https://www.wordandphrase.info/analyzeText.asp

### OUTCOMES AND BENEFITS

The immediate outcomes of our efforts include a set of 25 documents that are fully accessible to individuals with intellectual and developmental disabilities and others who read and listen with comprehension below a third-grade level. In the long term, we anticipate that our efforts to create the MTC Guidelines will lead to several new lines of inquiry and product development that will have positive outcomes and benefits for the target audience. These include (a) open-source tools for supporting stakeholders in producing texts that are consistently compliant with the MTC guidelines; (b) research regarding the ways that people with intellectual and developmental disabilities interact with and use texts created using the MTC Guidelines; and (c) research regarding the ways that features of text complexity might be manipulated in order to better support comprehension and literacy learning for children and adults with various profiles and needs. By increasing capacity to produce texts following the MTC Guidelines, while conducting investigations of useability, we hope that in the long term these guidelines increase opportunities for individuals with intellectual and developmental disabilities and others with extremely low literacy to independently access a myriad of texts regarding public health, emergency response, and emergency preparedness. This would have implications for spreading information, facilitating learning, increasing autonomy in decision making, and improving health outcomes for the target population.
DISCUSSION

Difficulties reading or listening to text with comprehension contribute to significant disparities in access to public health information and resulting health outcomes. With empirical evidence that adults with intellectual and developmental disabilities are up to eight times more likely to die after contracting COVID-19 (Gleason et al., 2021; Rabin, 2020), providing accessible information about protecting oneself and others remains of paramount importance. The need for this accessible information was recognized by the CDC and the CDC-F in the spring of 2020, early in the COVID-19 pandemic. The organizations responding by contracting with the Center for Inclusive Design and Innovation and the CLDS to create accessible versions of the CDC COVID-19 guidance documents. After the rapid review of the literature described here failed to reveal existing guidelines that could support the development of documents that would be accessible to people with intellectual and developmental disabilities who read and listen with comprehension at a third-grade level or lower, a new set of guidelines was created. The resulting research-based MTC Guidelines extracted information from the literature that was reviewed during the rapid review.

When the MTC Guidelines were applied to the original CDC COVID-19 guidance documents, the resulting documents were far less complex across all metrics. The simplified documents are significantly shorter, include sentences that are less complex and use more high-frequency words, and demonstrate high levels of consistency as demonstrated by comparison with the original documents and the standard deviations across metrics. Although the metrics suggest that the simplified documents will be easier to read, there is a need to evaluate the extent to which these documents can be read and understood by adults with intellectual and development disabilities who read at or below a third-grade level. Research is also needed to investigate whether reading these documents helps adults in the target population adopt practices to stay safe from COVID-19. Similarly, future research should also investigate how the MTC Guidelines might be used to improve access to other health and disaster preparedness information that is relevant to the target population.

We are encouraged by the success of the MTC guidelines to produce consistently less complex text because perceived difficulty has been shown to impact whether or not a text is read (Velayo, 1993). Both the metrics and side-by-side visual comparison of the two sets of documents suggest that the MTC documents will be perceived as easier to read. However, future research is needed to understand how the target population perceives the difficulty of the guidance documents. We are hopeful that people will return to the CDC for guidance in the future if they are able to have consistent and repeated opportunities to retrieve texts that they can read with comprehension. This is important because previous experience with, exposure to, and awareness of information resources has been shown to improve knowledge uptake (Buljan et al., 2018). Furthermore, repeated reading of texts with high internal consistency likely increases comprehension for the target population (Allington et al., 2015). Therefore, future research should investigate whether individuals are more likely to return to the CDC guidance once they are aware of the simplified resources and have success reading COVID-19 documents.
While the final documents are highly consistent and compliant with the MTC guidelines, we did not achieve this consistency immediately. Rather, just as the literature predicted (Hadden, 2015), it took many passes back and forth between authors in order to achieve the best iterations of texts possible. The time and degree of familiarity with the guidelines that was required to successfully create the set of COVID-19 documents described here limits the scalability of the approach. As such, future research should investigate how to operationalize where possible and scale the use of the MTC Guidelines, including training protocols and materials for authors to support integration of the MTC guidelines and resulting simplified documents into the production of written materials. Furthermore, it should be explored whether the MTC Guidelines are sufficiently detailed for the creation of smart automation tools to support writers in efficiently producing texts and checking compliance with the guidelines. Pending the outcomes of these areas of future research and development, effective application of the MTC Guidelines offers potential to increase personal access to vital information across a broad range of life sectors for adults with intellectual and developmental disabilities and others who read and listen with comprehension at a third-grade level and lower.

DECLARATIONS

The findings and conclusions in this report are those of the author(s) and do not necessarily represent the official position of the Centers for Disease Control and Prevention or ATIA. Development of these materials was supported in part by a grant from the CDC Foundation, using funding provided by its donors. The materials were created by the Center for Literacy & Disability Studies, Department of Allied Health Sciences, University of North Carolina at Chapel Hill, and Center for Inclusive Design & Innovation (CIDI), Georgia Tech. The CDC Foundation and Centers for Disease Control and Prevention (CDC) provided subject matter expertise and approved the content. The use of the names of private entities, products, or enterprises is for identification purposes only and does not imply CDC Foundation or CDC endorsement.

REFERENCES


