

## **Examining Readability in Regulatory Comments: A Digital Chasm**

### **Abstract:**

Theories of regulation and political communication assume that citizens are competently able to express their preferences to regulators. In internet mediated venues, communicative competence entails writing in a shared language that abides by shared conventions such as topical relevance, clarity and grammar. One measure of effective communication is the readability of an utterance, such as a comment made to regulators via regulations.gov. In this paper, we compare the readability of comments submitted to five heavily commented upon dockets in regulations.gov according to five measures: Fleish-Kincaid Reading Ease, Fleish-Kincaid Grade Level, Gunning Fog Index, McLaughlin's Simple Measure of Gobbledygook, and the Average Grade Level. We find that there are meaningful differences between the readability of the docket language and the comments made to those dockets. We also find meaningful differences in readability scores across the various dockets.

### **Keywords:**

Regulations.gov, communication, readability, responsive regulation

**Manuscript words:** 7937

**Tables:** 9

**Figures:** 6

## **Examining Readability in Regulatory Comments: A Digital Chasm**

Theorists of responsive regulation propose that regulation will be most effective when regulators and those they regulate are able to effectively communicate with one another (Baldwin and Black 2008). Within the well-known theory of responsive regulation proposed by Ayers and Braithwaite (1992), the key motivator to move up or down the pyramids of supports or pyramids of sanctions is open communication and dialogue. In his distillation of the “nine principles of responsive regulation”, Braithwaite makes it clear that communication is an essential component of each principle (Braithwaite 2011). In this article, we were concerned to look, at a basic level, at whether some basic conditions for responsive communication between regulators and the regulated were being met in the e-rulemaking space, regulations.gov. In particular we were curious to see whether the conditions for regulators and regulatees to “listen actively”, “give voice”, “build commitment” and “communicate lessons learned” were being met.

Responsive regulation is a theory of effective communication in a political setting, but theorists of political communication make it clear that specifically political communication, when effective, has a number of pre-conditions (see Habermas (1970), McCarthy (1973), Young (1997, 2002), Guttman and Thompson (2002), Taylor (1997) and Petraglia 2009). To communicate requires, at least, use of a shared language with shared rules. Other requirements, such as those for effective deliberation, such as described in Guttman and Thompson’s deliberation, Young’s inclusive politics, Taylor’s recognition or Habermas’ communicative action, depend first on the use of language and grammar by both parties in ways that both sides

of the communicative situation can understand. To empirical scholars of regulatory politics, the theoretical requirements of basic shared language for effective regulatory communication may seem obvious. However, the extent to which language used in the venue of citizen regulatory comment platforms meets this requirement is a less obvious matter for empirical assessment.

The purpose of this article is to report the findings of a research project addressing the readability of the language that commenters used when addressing the five most commented upon regulatory docket items in 2016. Regardless of the problematics of political communication, when it comes to written forms of political expression, such as making comments on docket items via Regulations.gov, the matter of comprehensibility looms large. In this article, we estimate how large the problem of making oneself understood is within the regulatory space. To do so, we analyze a sample of comments from five rulemaking exercises to determine the extent of the “readability” problem. Within the ambit of the original project, we were struck by the presentation of some commenters notes to regulators. Some were grammatically correct, well-reasoned and credible. Others demonstrated a lack of basic message design and grammar. While some comments were readable only to a highly sophisticated audience, others were not readable to any audience. At their core, the comments speak more to a breakdown in communication.

In understanding the extent of the problem, we seek insight by comparing the readability of docket items and comments made via Regulations.gov. Specifically, there are a number of concepts we explore:

- The readability of the docket items;

- The relationship between readability of the proposed docket and public comments; and
- The readability as a characteristic of the citizen participant.

We believe that in studying readability, we can draw out a number of worthy discussion points.

### **Informal Rulemaking and Regulations.gov**

As is well known to scholars of regulations, the 1946 Administrative Procedures Act (5 U.S.C. § 553) set out the minimal conditions for agencies to engage the public in informal rulemaking. Paramount among the requirements is that agencies publish information to the public about proceedings, including the text of the proposed rules, and that agencies provide the public opportunities to address the proposed rule “through submission of written data, views, or arguments with or without the opportunity for oral presentation” (§ 553(c)). The opportunity for public comment provides a channel for democratic control of rulemaking processes (though see McCubbins, Noll and Weingast (1987) and Balla (1998) on the APA as an instrument of political control).

Over the decades of enforcement of the APA, the evolution of communication technology has made commenting on rules more readily available. For example, the publication of the Federal Register via the internet (starting in 1994) has moved rulemaking from the Federal Depository Libraries to the desktop or laptops of interested persons around the world. Likewise, the transition of the Federal Register from a printed version to an on-line version has eliminated the expense associated with subscribing to the Federal Register and to storing it. Moving the

process of commenting upon rules to regulations.gov in 2003 also eliminated the costs of submitting letters or making phone calls to the agencies proposing rules.

Regulations.gov provides a low-barriers to entry framework where both regular and irregular participants have a common platform for communication in regulatory politics (Gormley 1986). The effective elimination of barriers to participation in rulemaking has prompted some scholars, such as Bryer (2013) to question whether Regulations.gov has actually made participation in the regulatory process too easy. Schulman (2009) has suggested that the ease of regulations.gov makes mass-mail or plebiscitary participation more available than ever. Others, like Holzer (2004, p. 15), claim “Digital deliberation broadens participation in the policy process”. Whether the openness of regulations.gov is helpful or hurtful to democratic deliberation is a higher order question than what we address in the remainder of this paper. Specifically, we ask at what level are the comments submitted through regulations.gov readable and comprehensible?

### **Plain Language and the Federal Government**

Placing pathologies of political communication at the feet of citizens alone would be a mistake. Problems of establishing effective communication with the citizenry is not new. There have been a number of initiatives over the past 15 years dedicated to improving communication, of which making governmental information web-available and free is just one component part. Suggested improvements to government to citizen communication have focused on two distinct areas: language comprehension and accessibility. The first is captured in the Plain Writing Act of 2010. The stated purpose of this act was “to improve the effectiveness and accountability of

Federal agencies to the public by promoting clear Government communication that the public can understand and use” (PLaw 11-274, section 2). Regulation is explicitly excluded from the Act. The Act specifies that “plain writing” or “writing that is clear, concise, well-organized, and follows other best practices appropriate to the subject or field and intended audience” must be used in three specific “covered document” types:

- “A document that is necessary for obtaining any federal government benefit or service, or filing taxes;
- Provides information about a federal government benefit or service; or
- Explains to the public how to comply with a requirement that the federal government administers or enforces”.

While this act made it clear that all Federal agencies should endeavor to make themselves understood to interested citizens, Executive Order 13563 (18 January 2011) made it clear that regulatory agencies were to improve access to regulations and regulatory review. This Executive Order augmented the terms of the plain writing act by stipulating that the regulatory system “must ensure that regulations are accessible, consistent, written in plain language and easy to understand”, while also being “adopted through a process that involves public participation”.

Public participation via the internet was specifically singled out in this Executive Order:

then-President Barak Obama ordered federal agencies to:

afford the public a meaningful opportunity to comment through the Internet on any proposed regulation, with a comment period that should generally be at least 60 days. To the extent feasible and permitted by law, each agency shall also provide for both proposed and final rules, timely online access to the rulemaking docket on

regulations.gov, including relevant scientific and technical findings, in an open format that can be easily searched and downloaded (Section 2b).

This executive order also mandated that the scientific evidence presented for review be both available for comment and conform to the standards for “Scientific Integrity” outlined in the President’s Memorandum for the Heads of Executive Departments and Agencies (March 9, 2009) (section 5).

The drive to create plain language in government has two primary drivers, one associated with basic communication needs and the other associated with the needs of a deliberative democracy. We review these in the next section below.

### **Literature Review**

One basic tenet of communication is that the writer must understand the audience. Kozol (1986, p.1-2) found that 60 million American adults, or one-third of the adult population, were reading at or below the eighth-grade level. In 2002, the National Adult Literacy Study found that 46 percent of adults, or over 90 million people, only functioned at the lowest two levels of literacy and comprehension. (Kirsch, Jungeblut, Jenkins & Kolstad, p. xvi). The concept of “functional illiteracy” is shifting, based on how one defines functionality; we agree with Kozol in that there is a significant difference between reading a fast-food menu and being an informed participant in the democratic process. The reading level of citizens who might comment upon regulations has a notable effect on how agencies must present their information to attain the democratic impulse instantiated into plain language rules. To wit, unless a regulatory docket is readable for the intended audience, it is unlikely that meaningful, much less responsive, regulatory participation will occur.

In contradistinction to the work done by Theil (2011), who studied an extensive number of elements of government websites and created a valuable set of benchmarks for optimal content, we consider readability as a barrier to effective interaction. Much of the literature discussing the intersection of readability and government focuses on the readability of government websites. Ferguson (2014) performed one of the more comprehensive analyses of municipal websites and found that the majority are written at a higher grade level than the target audience. Ferguson also identified that there were limited alternatives to text-based communication for those reading at a level below that of the website. Ferguson performs a thorough analysis of readability scoring methods, landing on Flesch-Kincaid Reading Ease and Flesch-Kincaid Grade Level as calculated by Microsoft Word as the measure for websites. A sample of municipal web sites from cities with more than 5,000 cities were assessed. Flesch-Kincaid Reading Ease for those sites as calculated as having a mean of 38.71, a median of 39.4 and a mode of 40. Flesch-Kincaid Grade Level was calculated as having a mean of 12.7, a median of 12.5, and a mode of 13.3. In short, Ferguson identified that most municipal web sites require the user to have at least a high school diploma to be able to understand the content. (Ferguson, 2014, p. 98).

Ferguson's work is given weight when considered alongside reports such as Visible Thread's *UK Local Authorities Website Clarity Index II*. (2016). Visible Thread's report is more of a meta-analysis designed to promote compliance with the United Kingdom's online plain language initiative. Visible Thread's analysis shows that only 18 percent of UK local government sites were at a Flesch-Kincaid Reading Ease score of 60 or higher, equivalent to an 8<sup>th</sup>-9<sup>th</sup> grade education. The range for Reading Ease in *Clarity Index II* is 46 – 90, which is both

narrower and above Ferguson's range of 0 to 74. (Ferguson, 2014, p. 98). In seeking to identify a standard for readability, we take note that Jonathan Kozol's 1986 report *Where stands the Republic? Illiteracy: a warning and a challenge to the press* is often cited. We were only able to locate a single copy of the source document, which was obtained via interlibrary loan from Georgia State University.

The most parallel work involves a readability assessment of the ClinicalTrials.gov site. (Wu, et al., 2015). ClinicalTrials.gov serves as a registry for clinical trials related to medical interventions. The site specifically includes "patients, their family members... and the public" as part of the target audience. (U.S. National Institutes of Health, 2016). ClinicalTrials.gov is legally required to provide such information in an easy-to-understand form. (Wu, et al., 2015), (FDA Modernization Act of 1997). The authors found that clinical trial information is being presented in a manner that required 18 years of education on average to read effectively. The study found a critical need to improve readability of clinical trial information. (Wu, et al., 2015).

The readability of government websites is important for normative, democratic, reasons. Without the ability to read and respond to the information, citizens cannot meaningfully participate in the governance of their lives and businesses. Low mechanical barriers to entry, such as publication in the free venue of the internet helps to create pre-conditions for more participation, but it is not implied that more is better participation.

Holzer (2004) characterizes Regulations.gov specifically as a type of information and communication technology to facilitate digital citizen, i.e. democratic, participation. Holzer classifies communication as static or dynamic participation, creating conditions for either information dissemination or citizen deliberation. To that end, Regulations.gov is a static

platform for regulators to provide information to the citizenry, and for the citizenry to provide feedback. Communication via Regulations.gov is simple in nature and there is no opportunity for true deliberation. London (1995) describes initiatives such as Regulations.gov as teledemocracy; in this model viewpoints are aggregated and congealed into a collective choice in a fair and efficient manner. This does not imply, however, that the aggregation of viewpoints yields anything substantively important for the attainment of democratic deliberation.

### **Readability Measures**

Both Ley and Florio (1996) and Ferguson (2014) provide in-depth discussions of the various readability measures. Readability measures in general utilize a number of factors, such as number of words, number of sentences, average words per sentence, average syllables per word, proportion of words which are either monosyllabic or more than three syllables, and proportion of common words. The readability measures we elect to use in our project include the Flesch-Kincaid Reading Ease (FKRE), the Flesch-Kincaid Grade Level (FKGL), McLaughlin's SMOG Grading (SMOG), the Gunning Fog Index (GFI), and an average grade index returned by the measurement tool.

Flesch-Kincaid Reading Ease is a standardized scoring algorithm that returns a numerical score of 0 to 100, with lower scores correlating to increased reading difficulty. The FKRE is used extensively throughout the U.S. government. As we will discuss later, we determined that there are variations among the automated engines used to generate the specific grade-level readability scores. We elected to use FKRE as our primary measure because of the potential

range. The FKRE relies on the number of words, number of sentences, and number of syllables in computing readability.

The FKRE is easily translated to an approximate grade level (FKGL). Translating the FKRE to an approximate grade level is useful when comparing results with other readability measures, or communicating the results of readability scoring to others. *Clarity Index II* (Visible Thread, 2016) utilizes a score of 60 as the threshold for success with U.K. municipal websites. A score of 60 is established as a “standard” reading level and approximates to the reading abilities of a person who has a 9<sup>th</sup>-10<sup>th</sup> grade education.

McLaughlin’s SMOG scoring method relies on an analysis of polysyllabic words in a passage. Unlike the FKRE, the SMOG score is based on an analysis of 30 actual or constructed sentences, rather than word and sentence counts. The Gunning Fog Index (GFOG) is based on an analysis of complex words of more than three syllables within the context of average sentence length. GFOG disregards proper nouns, jargon and compound words, as well as common suffixes. One of the issues associated with GFOG is that not all words of three or more syllables are difficult to read; words such as “America” are easily recognized and understood, while some monosyllabic words such as “faux” are more difficult and less commonly understood. We include the GFOG and SMOG scores as comparisons with the FKRE and FKGL. We also include an average grade level, which is computed by the readability service and represents the mean of the seven separate readability scores calculated.

## **Methods**

Regulations.Gov allows programmatic retrieval of documents and information via an application programming interface (API), opened to the public in 2011. This API is extremely useful but has distinct limitations. First, in order to use the Regulations.Gov API, a key must be requested. Second, the API places limits requests to 1000 records per call, and users may only make 1000 calls to the database per hour.

We selected the five most commented upon regulations as described by regulations.gov as our sample of regulatory comments. While this may have produced some sampling error-- all of the comments drew a large number of persons which may not represent the modal commenter-- we selected this to generate a sufficiently large body of text for analysis. Once the specific dockets to be analyzed were selected, a method of retrieving and storing the public submissions via API was devised. This entailed two distinct components.

Two distinct components were necessary to retrieve the data. First, a default function was written to retrieve the contents via a URL request and return the results to the script. The second component was the script to make the appropriate request to Regulations.Gov and store the resultant data. The function and script were written in PHP, a server-side scripting language already in use by the university system of the authors. Data was requested in extensible markup language (XML) format for easy conversion and analysis using conventional tools such as Microsoft Excel. The method used does not generate an XML schema with each file, only the data.

The API Key was used to return general information about the dockets via the Regulations.Gov Developer web site (<http://regulationsgov.github.io/developers/console/>). The number of comments was noted for use in the script. The script initiates by defining three

variables, including the docket number as `$docket`, the number of comments as `$docketUL` (docket upper limit), and the number of requests necessary to retrieve all comments as `$docketPasses`. The latter was calculated by dividing the number of comments by the number of comments returned per request, and adding one additional to ensure capture of all comments. The base URL is constructed using the data items required by Regulations.Gov, including the API key, `countsOnly` set to 0 so that documents are returned, `encoded` set to 1 so that only Regulations.Gov parameters are accepted, the document type set to public submissions, the docket identification string, `results per page` set to the maximum of 1000, and a page offset. The script sets the variables for page offset (`$offset`) and pass (`$pass`) to zero.

A programmatic loop is called using a comparison between the current pass through the loop, and the number of passes determined as necessary in the variable `$docketPasses`. The variable `$uri` is created by combining the base URL and current page offset. The function to retrieve the information from the URL is called and the data stored in a variable called `$sXML`. The variable `$logFileName` is created using a text string. With each docket queried, the string was modified to reflect the specific docket. In the example code, this is represented as `<1>`. Information contained in the variable `$pass` is appended so that each data set within the loop is easily identified, as is the extension “.xml”. The contents of the variable `$sXML` is appended to the file. The variable `$pass` is then incremented by one, and the variable `$offset` is incremented by 1000. The loop is closed.

When run, the script loops through the requisite number of passes, retrieves the data set, and creates an XML file for each 1000 documents. The XML files are in the same directory as the page running the script. For extremely large data sets, the server running the script may

experience timeout issues. If so, then the PHP option `set_time_limit` may need to be increased beyond the default 30 seconds.

The script was run on each of the five dockets studied. As each script completed, the resultant files were moved to a read-only directory and stored. A master XML file was created for each Docket then a working XML file was created to handle the XML import process. In order to facilitate analysis of the text of the submitted comments, the text comments from each docket's master XML file was placed in a comma-separated values (CSV) file. The Readability-score.com engine requires that files be less than five megabytes in size to be scored properly. The comments CSV file for each individual docket was opened and saved in a new CSV files containing 5000 comments. Each file was submitted to Readability-score.com for measurement. The resultant CSV files were combined using the Windows command line. Within the directory, "Copy \*.csv Docket<#>.csv" was run, with the # changed based on the docket being combined. Extraneous header lines, blanks, and lines that returned errors were removed in Excel.

### **Data**

Readability scores typically require passages to contain 150 words or more. (Ferguson, 2014, p. 102). The output from Readability-score.com includes a word count as a unique column, so a filter was applied to exclude all rows with a word count of less than 150. The compiled score file was then saved as a Microsoft Excel workbook. Additionally, the scoring engine typically returned a small number of values (typically  $n < 5$ ) within each file that were outside the FKRE's range of 0 – 100. Upon evaluation, these were often nonsensical comments, or simply repeated monosyllabic words. Comments that resulted in a score outside the FKRE range

were excluded from further analysis. Table 1 presents the total number of comments in each docket along with those analysed based upon the sampling procedures described.

<Insert Table 1 about here>

The raw readability score file consists of eight individual readability scores, an average grade level score based on the individual scores, counts of sentences, letters, words, and syllables. Additional calculated fields are provided, as well as basic sentiment and keyword analysis. As discussed earlier, we elected to use the FKRE, FKGL, GFOG, SMOG, and AGL in this project. For each readability score of interest, the statistical mean, median, mode and standard deviation were calculated using Microsoft Excel's native functions. The mean values are represented in Table 2 and Figure 1 below.

<Insert Table 2 about here>

<Insert Figure 1 about here>

Based upon the scores represented here, it seems that the efforts of the plain writing in regulations acts and orders were not particularly successful at bringing the level of language in the rulemaking dockets down to the level of citizens. In the case of the FKRE, a lower score indicates greater level of difficulty to read. In all of the cases, the score of the rule was lower than that of the comments. In the case of the reading level, there was variation in the estimated grade level necessary to comprehend the comments versus the rule. For the third and fifth rule, the comments were more difficult to read than the rule itself. However, according to the

Gunning Fog index score, which offers a measure of the number of years of education a reader would need to have to grasp the content, all rules except for that in docket three (Department of State) require more years to understand than the comments attached to that rule. The same pattern is seen for the SMOG index score and the average grade level scores. This will be examined in greater detail in the analysis of each docket below.

**Docket 1: CFPB-2016-0025. Docket 1.** Docket CFPB-2016-0025 is a proposal to regulate short term, small amount lending commonly referred to as “payday loans”. Payday lending is a significant financial service in the United States, with 12 million Americans per year utilizing this type of financial service. (Pew Charitable Trusts, 2012, p. 4). Using Gormley’s typology, this docket item was identified as low in complexity salience and exemplar of street-level politics. The readability scoring of both the proposed action and public comments are very much in line with what the expectation.

The proposed regulatory action sets the operational playing field for the three major stakeholder groups - government, the payday loan providers, and payday loan consumers. The docket lies directly within Gormley’s street-level framework. At over 420,000 words, the proposed action is around 60 times longer than the other docket items evaluated in this project. The proposed action provides operations-level detail, which correlates directly with Gormley’s description of street-level politics as having a “set of routines that afford predictability.” (Gormley, Jr., 1986, p. 610). We expected there to be a large number of comments, primarily from organizations and consumers impacted by the proposed action. As outlined in our other

research, this expectation was largely met by the data. From that initial footing, we are led to look deeper into the proposed action and public comments themselves.

Readability scoring reveals that the proposed action requires reading skills well above that found in the payday lending consumer. The text of the proposed regulation has a FKRE score of 27.8, which is extremely difficult to read. The FKGL is 17, equating to the reading ability of a first-year graduate student. The public comments were normally distributed and have a mean FKRE of 74.1 and mean FKGL of 7.5, which is consistent with the payday lending consumer. Only 14 percent of payday loan users identified as having a college degree, and 16 percent identified as having less than a high-school education. (Pew Charitable Trusts, 2012, p. 35). Docket 1 was the only item studied that resulted in a mean FKRE above 60 on analyzed comments. In fact, over 99 percent of analyzed comments (n = 70,301) had an FKRE above 60.

Docket 1 also presented a unique characteristic in how widely the mean readability of the comments differed from the readability of the docket item. Even though the number of comments differed significantly, a fairly consistent pattern emerged with Dockets 2-4. We suspect that this results from activation of the payday-lending customer base by the payday lenders. Summary statistics and distribution of scores are presented in Table 3 and Figure 2 below.

<Insert Table 3 about here>

<Insert Figure 2 about here>

**Docket 2.** Docket VA-2016-VHA-0011 is a proposal to allow greater flexibility for advance practice nursing professionals to provide health care within the Veteran’s Administration health

care system. We classified this proposed action as low complexity, high salience under Gormley. This places the action within the sphere of “hearing room politics”. (Gormley, Jr., 1986, p. 607). In keeping with that, we expected public comments to have a normal distribution clustered around the standard FKRE of 60. We felt that given the high-profile media coverage of administrative and health care problems within the Veteran’s Administration, there would be a higher level of public comment and the FKRE normally distributed around the standard score of 60.

The proposed action’s readability was as expected, with a FKRE of 38.5 and FKGL of 12.4, making it difficult to read but comparatively easier than Docket 1. When the FKRE frequency distribution was graphed, there was an unexpected cluster of comments that scored as extremely hard to read. Nearly 10 percent (n = 1252) scored with an FKRE of 21 or below. Although there was an expected shift in readability, there was also a left skew. Very few comments scored above the FKRE standard of 60 (n = 568).

Further investigation revealed two distinct patterns. Comments that scored extremely difficult to read were largely from radiologic technologists, radiologists, and similar health care professionals that would potentially find their roles within the system changed if the proposed action moved forward. The comments that scored above FKRE 60 appeared to be from veterans receiving healthcare from the Veteran’s Administration system and their family members. Table 4 and Figure 3 offer statistics on the comments for this docket.

<Insert Table 4 about here>

<Insert Figure 3 about here>

**Docket 3.** As previously noted, Docket 3 yielded 7,665 comments that were successfully analyzed by the readability scoring engine. The low number of comments analysed in this action results from the high number of duplicate comments (over 100,000) submitted under this docket heading.

Docket DOS-2014-0003 is different from other rulemaking exercises in this exercises: it is a “notice of a 30 day comment period regarding the national interest determination for the TransCanada Keystone Pipeline”. Although this is not a rulemaking exercise the notice action asks for citizens to comment upon the Presidential Permit to “authorise construction, connection, operation and maintenance” of the pipeline. As listed in the notice, “The border facilities would be part of a proposed 875-mile pipeline and related facilities (the Keystone XL project) that is designed to transport up to 830,000 barrels per day of crude oil from Alberta, Canada and the Bakken shale formation in North Dakota and Montana. The pipeline would cross the U.S. border near Morgan, Montana and continue through Montana, South Dakota, and Nebraska, where it would connect to existing pipeline facilities near Steele City, Nebraska for onward delivery to Cushing, Oklahoma and the U.S. Gulf Coast region”.

We classified this proposed action as high complexity and high salience under Gormley, but its unique high visibility added a new dimension to the urgency of public comment. Due to the salience, complexity, and visibility of the project we expected public comments to have a non-normal distribution with comment text readability bi-modally distributed with some comments being exceptionally sophisticated and others being less so.

The proposed action’s readability was not as expected, with a FKRE of 54.25 and FKGL of 9.79, making it relatively easy to read. When the FKRE frequency distribution was graphed,

there was an apparent cluster around the mean score of 50, with a large number of comments on either side, but more between a score of 55 and 60. Contrary to our expectations of a bimodal distribution, there was a somewhat normal distribution to the comments with very few comments being exceptionally simple or exceptionally complex.

As mentioned, a tremendous number of duplicate comments were culled from this docket. Unlike in other dockets, these were exact duplicates submitted by individuals with the same organizational affiliation--BKM Strategies. No individual modification of this comment text was made, as we found in other dockets. Other unusual patterns were that individuals “called out” other commenters as being “Industry insiders” or reproduced one another’s comments with the suggestion that the original comment was fraudulent due to an affiliation of the original commenter with an energy company. These exceptional patterns, represented in Table 5 and Figure 4 below, may have skewed the findings from this otherwise unusual component of the rulemaking processes under review here.

**Docket 4.** As previously noted, 11,376 comments from Docket 4 were successfully analyzed by the readability scoring engine. This rulemaking action was intended to “guide to tax-exempt social welfare organizations on political activities related to candidates that will not be considered to promote social welfare”. This rulemaking action was proposed to differentiate common good or general welfare activities of organizations from political activities. Perhaps due to the storm at the time around the head of the IRS Lois Lerner and regulation of PACs and political organizations with a right-leaning bent, the comments submitted to this rule were quite

varied, quote political and, in spots, quite surprising in the use of curse words, racial epithets, and accusations of anti-democratic conduct.

We categorized this rule as being low salience and low complexity. Despite this being an IRS rule, which could be expected to be high complexity due to its relationship to the tax-code, this rule addressed a low complexity topic of categorizing organizations. As this low complexity and low salience rule could not be expected to bring substantial interest group, politician or press interest, we expected that the comments would be quite varied in their readability as few would be directed by strong organizational forces. To that end, we expected no strong trends in the comment readability.

What we found was that there was a strong, downward, trend in the FKRE readability score for the comments (mean of 34.6), meaning that the comments were, overall, difficult to read. Other readability indicators also suggested that the comments were overall difficult to read. This is unusual given the colorful language found in the comment text; it strikes the present authors as unusual that the grade level for comment text that includes racial epithets (n\*\*\*\*\*), gendered slurs (c\*\*\*), and repeat use of curse words (e.g., f\*\*\*) would also be among the most grammatically difficult to read. Table 6 and Figure 5 present information about these comments.

<Insert Table 6 about here>

<Insert Figure 5 about here>

**Docket 5.** As with docket 3, docket 5 yielded relatively few (1,167) non-duplicate comments that were successfully analyzed by the readability scoring engine. This docket addressed a strongly controversial matter-- funding for reproductive services under Title X grants-- couched in highly technical language. Although the comment texts make repeated reference to abortion and other women's health and reproductive services, the docket title and summary do not suggest that this is part of the substance of the rule:

“This document seeks comment on the proposed amendment of Title X regulations specifying the requirements Title X projects must meet to be eligible for awards. The amendment precludes project recipients from using criteria in their selection of subrecipients that are unrelated to the ability to deliver services to program beneficiaries in an effective manner.”

The meaning of the rule is to change terms of participation of subrecipients of awards to states under the Title X Family Planning Program. Changes to terms of grant flow though are complex, but in this rulemaking description, the grant focused complexity is low. As the substantive changes would alter a highly salient area, we discussed this as an example of hearing room politics. Under these assumptions, we assumed the comment text would be quite easy to read and that the distribution of readability scores would cluster around a low grade level. We did not find that to be the case: the readability scores were close to the mean for all dockets and close to normally distributed. The higher grade level scores may be attributable to the high number of slightly transformed duplicates included in this readability analysis. In this docket, there were four standard form letters repeatedly submitted. These form letters were long, well composed, grammatically correct, and used sophisticated political language. We surmise that,

even when personalized, the readability on these letters would have trended towards higher grade levels.

<Insert Table 7 about here>

<Insert Figure 6 about here>

## **Discussion**

**Measurement of Readability.** As we were developing the methodology for this project, a number of readability scoring tools were evaluated. Scores were repeatable within each tool, but there was a wider variation between the individual tools than was expected. We could find no empirical studies comparing the accuracy of various automated readability scoring engines, and no studies comparing manual calculation of readability with automated readability calculation by software. Sirico (2008) states that readability scoring in consumer products such as Microsoft Word is “mechanical and imperfect,” yet remains popular in part because of availability and ease of use. Three automated scoring engines were identified initially:

- Flesh 2.0, an open-source tool that analyzes the readability of a file;
- Readability-Score.com, an online service; and
- Microsoft Word’s integrated analysis tool.

Sirico (2008) used a simple passage of 151 words in his work and found a difference of 0.84 grade levels between manual scoring and Microsoft Word. Sirico goes on to state “different

software packages that purportedly use the same formula sometimes yield differing results for the same textual sample.” Sirico also tested a number of online tools, which yielded a wide range of scores as well. We acknowledge and agree that differing tools yield a range of results.

We recognize that inconsistency in data is a major research concern, and as Siciro states “the lure of technology has reduced the accuracy of at least one test for readability.” To illustrate the variation, we elected to test the language of Docket 5, HHS-OS-2016-0014 using multiple tools. The two software-based tools scored the sample as more difficult to read than the online tools. The results of this comparison are listed in Table 8 below.

<Insert Table 8 about here>

Ultimately, we elected to use Readability-Score.com as our principal tool for readability scoring. Our rationale to do so was based on several factors:

- The ability to score individual comments individually, rather than en masse. Sicaro (2008) points out that in lengthy documents, individual paragraphs or sections may yield significantly different scores;
- The ability to score a much higher volume of text than studied by Bryer (2013) in an automated, consistent fashion; and
- Readability-Score.com has previously been used in scholarly work.

**Readability of Proposed Actions.** One of the core findings of the project is that the individual docket items presented for comment are well above the reading level of most citizens. We agree

with other researchers that readability remains an issue for many governmental and governmentally-required communications despite the numerous efforts at improving readability. Our findings mirror similar research, including both Ley and Florio (1996, p. 7) with regard to important health care information and Wu, *et al* (2015, p. 274) with regard to clinical trials.

As previously discussed a target FKRE of 60 is a de facto standard for public communications. None of the docket items reached that target, with four of the items scoring below 40 on the FKRE and requiring at least a 12<sup>th</sup> grade education by score. Docket 3, DOS-2014-0003, the sole exception, scored as the most readable at 51.1 FKRE and 8.6 FKGL. Docket 3 is unique in that the primary document is not a proposed action; the primary document is a brief outline of the status of the overarching project with a request for public comment. Docket 3's primary document has only 538 words and was the shortest of the five analyzed. These scores are presented in Table 9 below.

<Insert Table 9 about here>

We noted that with regard to docket items, the lowest FKRE corresponded with the highest word count, and the highest FKRE with the lowest word count. Calculating the Pearson Correlation Coefficient between FKRE and word count yielded a score of -0.5947, a moderately negative correlation. We acknowledge that the docket content found on Regulations.gov is unique, in that the items contain short titles, definitions and headings which could impact readability scoring. This uniqueness does not sway the basic underlying problem – proposed

government regulations and actions range from very difficult to nearly impossible for the average American to read.

In an article by McGrane (2009), Rutgers' Ross Baker is quoted as saying "These bills are not written for even the educated layperson. They are written for specialists." As noted by Coglianese (2006, p. 967), there is little hope that technology alone can overcome "the deep motivational, cognitive, and knowledge-based chasms that stand in the way of citizen participation in the regulatory process." The readability findings as related to the docket items themselves lend support these broad generalizations.

We also evaluated the relationship between FKRE of the docket item and the volume of freeform text public comments analyzed. Surprisingly, the least-readable docket item had the largest number of comments to analyze. There was a moderately negative Pearson Correlation Coefficient of -0.6102. We find this relationship interesting and note for future study. Bryer (2013) specifically noted relevance but an absence of logic and objectivity in public commenting. We question whether the absence of that logic and objectivity is related to a failure by the public to read and understand the docket item.

**Power.** We are struck by the power dynamic presented in the data, especially as related to Docket 1, Payday Lending. The proposed regulatory action is consistent with elite theory, given the length, readability, and purpose - protecting the common citizen. (McBeth, & Clemons, 2009). The proposed action is new to the federal government; prior consumer protections were enacted by individual states. In that regard, the decision to act is an exercise of power over both payday lending services and the public who utilizes them. In stark opposition to the proposed

regulation are the payday lending customers. As we reviewed the docket and comments, we discussed the motivation for so many individuals to be activated. The readability scoring reflected a much different commenter than the other four dockets studied, with a mean FKRE over 20 points higher and two grade levels lower. We suspect the answer lies in the power payday lending providers have over their consumers. The relationship between power and salience as a dimension of citizen engagement and e-government merits future examination.

### **Conclusion**

As pointed to in the introduction and elsewhere, the concern for readability of regulatory documents and for the readability of citizens interaction with regulations speaks to the goal of creating a democratically responsive government. Theorists of responsive regulation and political communication each point to the importance of both sides listening to and understanding one another. Without the ability of both sides-- regulator and regulatee-- to speak a similar language with similar rules, it is unlikely that the act of regulation will advance “up the pyramid” of rewards and “down the pyramid” of sanctions. Instead, regulation will appear to be a non-responsive, coercive mechanism heaped upon a public with no recourse to challenge the terms of their coercion. Advancing beyond this pessimism involves changing the regulatory environment in many ways, of which improvements to docket readability is only one. Improving public education generally and enhancing citizen’s understanding of the informal rulemaking processes are another way to enhance communication.

## References

- Ayres, I., & Braithwaite, J. (1992). *Responsive regulation: Transcending the deregulation debate*. Oxford University Press on Demand.
- Baer, J., Kutner, M., Sabatini, J., & White, S. (2009). *Basic Reading Skills and the Literacy of America's Least Literate Adults*. Washington, DC: National Center for Education Statistics.
- Baldwin, R., & Black, J. (2008). Really responsive regulation. *The Modern Law Review*, 71(1), 59-94.
- Balla, S. J. (1998). Administrative procedures and political control of the bureaucracy. *American Political Science Review*, 92(03), 663-673.
- Braithwaite, J. (2011). The essence of responsive regulation. *UBCL Rev.*, 44, 475.
- Bryer, T. (2013). Public participation in regulatory decision-making. *Public Performance & Management Review*, 37(2), 263-279.
- McBeth, M. K., & Clemons, R. S. (2009). *Public policy praxis: a case approach for understanding policy and analysis*. New York: Pearson / Longman.
- McCarthy, T. A. (1973). A theory of communicative competence. *Philosophy of the Social Sciences*, 3(2), 135-156.
- Coglianesi, C. (2006). Participation in rulemaking: Past, present and future. *Duke Law Journal*,

55(5), 943-968.

McCubbins, M. D., Noll, R. G., & Weingast, B. R. (1987). Administrative procedures as instruments of political control. *Journal of Law, Economics, & Organization*, 3(2), 243-277.

Fabian, J. (2010, August 25). *Key Senate Democrat suggests that he didn't read entire healthcare reform bill*. Retrieved from The Hill.com:  
<http://thehill.com/blogs/blog-briefing-room/news/115749-sen-baucus-suggests-he-did-not-read-entire-health-bill>

FDA Modernization Act of 1997, Pub. L. 105-115, 111 Stat. 2296, codified as amended at 42 USC 282. (n.d.).

Ferguson, D. (2014, May). *Grade-Level Readability of Municipal Websites: Are They Creating Digital Inequalities of Opportunities that Perpetuate the Digital Divide*. Akron, OH: ProQuest.

Gormley, Jr., W. T. (1986, Summer). Regulatory Issue Networks in a Federal System. *Polity*, 18(4), 595-620.

Gutmann, A., & Thompson, D. (2002). Deliberative democracy beyond process. *Journal of Political Philosophy*, 10(2), 153-174.

Habermas, J. (1970). Towards a theory of communicative competence. *Inquiry*, 13(1-4), 360-375.

- Holzer, M. (2004). Restoring Trust in Government: The potential of digital citizen participation. In M. Holzer, M. Zhang, & K. Dong (Ed.), *Frontiers of Public Administration* (pp. 6-23). Beijing: National Center for Public Productivity, Rutgers.
- Kirsch, I., Jungeblut, A., Jenkins, L., Kolstad, A. (2002, April). Adult Literacy in America: A first look at the findings of the National Adult Literacy Survey. Washington, DC: U.S. Department of Education.
- Kozol, J. (1986). *Where Stands the Republic? Illiteracy: A warning and a challenge to the Nation's Press*. Atlanta, GA: Cox Enterprises. Retrieved from <https://eric.ed.gov/?id=ED281029>
- Ley, P., & Florio, T. (1996). The use of readability formulas in health care. *Psychology, Health & Medicine, 1*(1), 7-28. doi:10.1080/13548509608400003
- London, S. (1995). Teledemocracy vs. Deliberative Democracy: A comparative look at two models of public talk. *Journal of International Computing and Technology, 3*(2), 33-55.
- McGrane, V. (2009, September 8). *Read the bill? It might not help*. Retrieved from Politico.Com: <http://www.politico.com/story/2009/09/read-the-bill-it-might-not-help-026846>
- Petraglia, J. (2009). The importance of being authentic: Persuasion, narration, and dialogue in health communication and education. *Health communication, 24*(2), 176-185.
- Pew Charitable Trusts. (2012). *Payday lending in America: Who borrows, where they borrow, and why*. Washington, DC: Pew Charitable Trusts. Retrieved from

[http://www.pewtrusts.org/~media/legacy/  
uploadedfiles/pes\\_assets/2012/pewpaydaylendingreportpdf.pdf](http://www.pewtrusts.org/~media/legacy/uploadedfiles/pes_assets/2012/pewpaydaylendingreportpdf.pdf)

Sirico, Jr., L. J. (2008). Readability Studies: How Technocentrism can Compromise Research and Legal Determinations. *Villanova University School of Law Working Paper Series(104)*. The Berkeley Electronic Press. Retrieved from <http://law.bepress.com/villanovalwps/art104>

Taylor, C. (1997). The politics of recognition. *New contexts of Canadian criticism*, 98.

Taylor, G. (2012). Readability of OHS documents - A comparison of surface characteristics of OHS text between some languages. *Safety Science*, 50, 1627-1635.  
doi:10.1016/j.ssci.2012.01.016

Thiele, L. (2011). *Online Government: An analysis of municipal websites in the United States*. Proquest Dissertation & Theses Global.

U.S. National Institutes of Health. (2016, September). *ClinicalTrials.gov Background*. Retrieved Dec 20, 2016, from ClinicalTrials.gov:  
<https://clinicaltrials.gov/ct2/about-site/background>

Visible Thread. (2016). *UK Local Authorities Website Clarity Index II*. Cambridge, UK: Visible Thread, Inc.

Wu, D. T., Hanauer, D. A., Mei, Q., Clark, P. M., An, L. C., Proulx, J., . . . Zheng, K. (2015, Aug 11). Assessing the readability of ClinicalTrials.gov. *Journal of the American Medical Informatics Association*, 23(2), 269-275. doi:10.1093/jamia/ocv062

Young, Iris Marion. *Intersecting voices: Dilemmas of gender, political philosophy, and policy*.

Princeton University Press, 1997.

Young, I. M. (2002). *Inclusion and democracy*. Oxford University Press on Demand.

**Tables and Figures for “Examining Readability in Regulatory Comments: A Digital Chasm”**

**Table 1: Reduction of Comments for Readability Analysis**

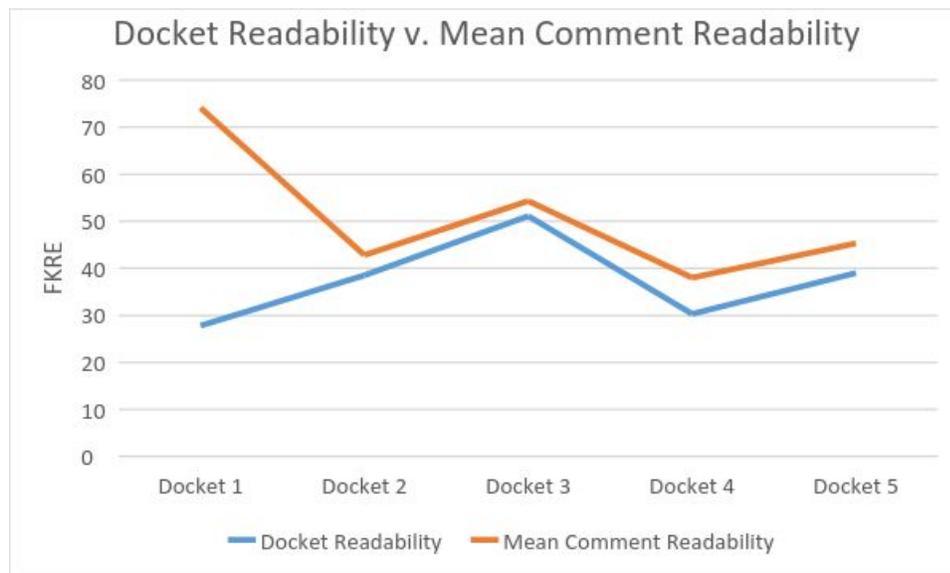
	<b>Docket 1 CFPB-201 6-0025</b>	<b>Docket 2 VA-2016-VHA -0011,</b>	<b>Docket 3 DOS-2014- 0003</b>	<b>Docket 4 IRS-2013-00 38</b>	<b>Docket 5 HHS-OS-201 6-0014</b>
<b>Retrieved Comments</b>	199,430	178,376	127,206	143,857	92,458
<b>Removal of Comments with Attachments</b>	-63,327	-57,904	-3,095	-1,069	-52,673
<b>Removal of Duplicate Comments</b>	-4,190	-85,929	-100,690	-79,856	-36,526
<b>Comments Submitted for Readability Scoring</b>	131,913	34,543	23,421	62,932	3,259
<b>Comments Returned</b>	129,867	34,543	23,421	62,926	3,259
<b>Removal of Comments &lt; 150 Words</b>	-58,876	-23,822	-15,766	-51,550	-2,092
<b>Comments Analyzed</b>	70,991	10,721	7,655	11,376	1,167

**Table 2: Mean of Readability Data by Docket**

	<b>Flesch-Kincaid Reading Ease</b>	<b>Flesch-Kincaid Grade Level</b>	<b>Gunning Fog</b>	<b>SMOG</b>	<b>Average Grade Level</b>
<b>Docket 1, CFPB-2016-0025, Payday Lending</b>					
<b>CFPB- 2016-0025 Rule Score</b>	27.76	17.04	**	**	**
<b>Comments Score (70,991)</b>	74.07	7.48	10.61	10.20	8.54
<b>Docket 2, VA-2016-VHA-0011, VA Advanced Practice Nursing</b>					

<b>VA-2016-VHA-001 1 Rule Score</b>	38.5	12.4	15.4	14.2	14.1
<b>Comments Score (10,721)</b>	42.84	11.83	14.64	13.61	13.43
<b>Docket 3, DOS-2014-0003, Keystone Pipeline</b>					
<b>DOS-2014-0003 Rule Score</b>	51.1	8.6	10.7	10.7	10.7
<b>Comments Score (7660)</b>	54.25	9.79	12.38	11.88	11.28
<b>Docket 4, IRS-2013-0038, IRS</b>					
<b>IRS-2013-0038 Rule Score</b>	30.3	14.2	18.1	16.1	15.9
<b>Comments Score (11,357)</b>	38.00	13.01	16.60	14.86	14.54
<b>Docket 5, HHS-OS-2016-0014, Title X Funding</b>					
<b>HHS-OS-2016-001 4 Rule Score</b>	39	12	16.1	14.6	14.1
<b>Comments Score (1167)</b>	45.29	12.08	15.21	13.94	13.77

**Figure i: Docket Readability Compared to Mean Comment Readability**

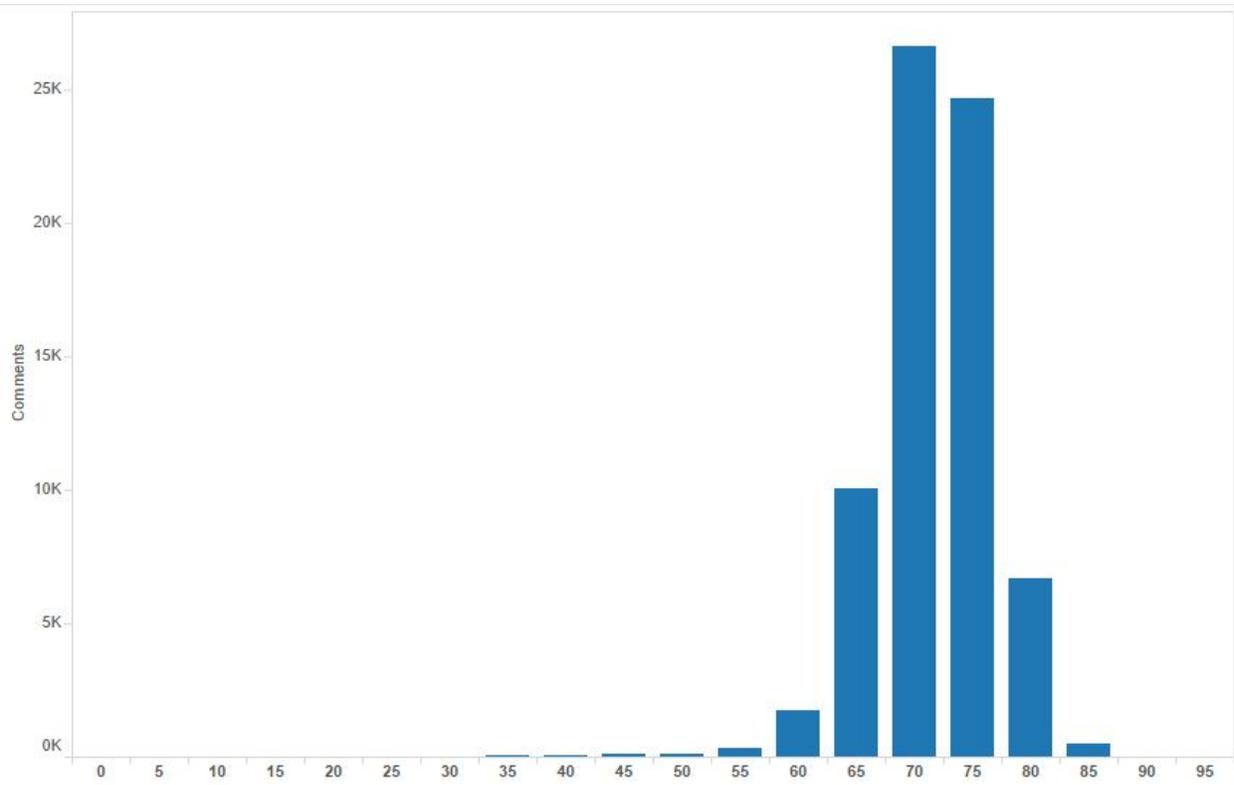


**Table 3: Readability scoring of public comments in response to Docket 1, CFPB-2016-0025.**

	<b>FKRE</b>	<b>FKGL</b>	<b>GFOG</b>	<b>SMOG</b>	<b>AGL</b>
--	-------------	-------------	-------------	-------------	------------

<b>Maximum Score</b>	96.9	36.5	40.5	17.8	29.2
<b>Minimum Score</b>	3.2	2.2	3.1	4.3	3.9
<b>Mean</b>	74.1	7.5	10.6	7.8	8.5
<b>Median</b>	74.3	7.4	10.5	7.7	8.5
<b>Mode</b>	73.5	7.4	10.4	7.7	8.4
<b>Standard Deviation</b>	5.16	1.08	1.22	1.01	0.99
<b>Within 1SD of Mean</b>	52964	51734	51127	51763	51407

**Figure 2: Comment Frequency Distribution by FKRE Score for Docket 1, CFPB-2016-0025.**

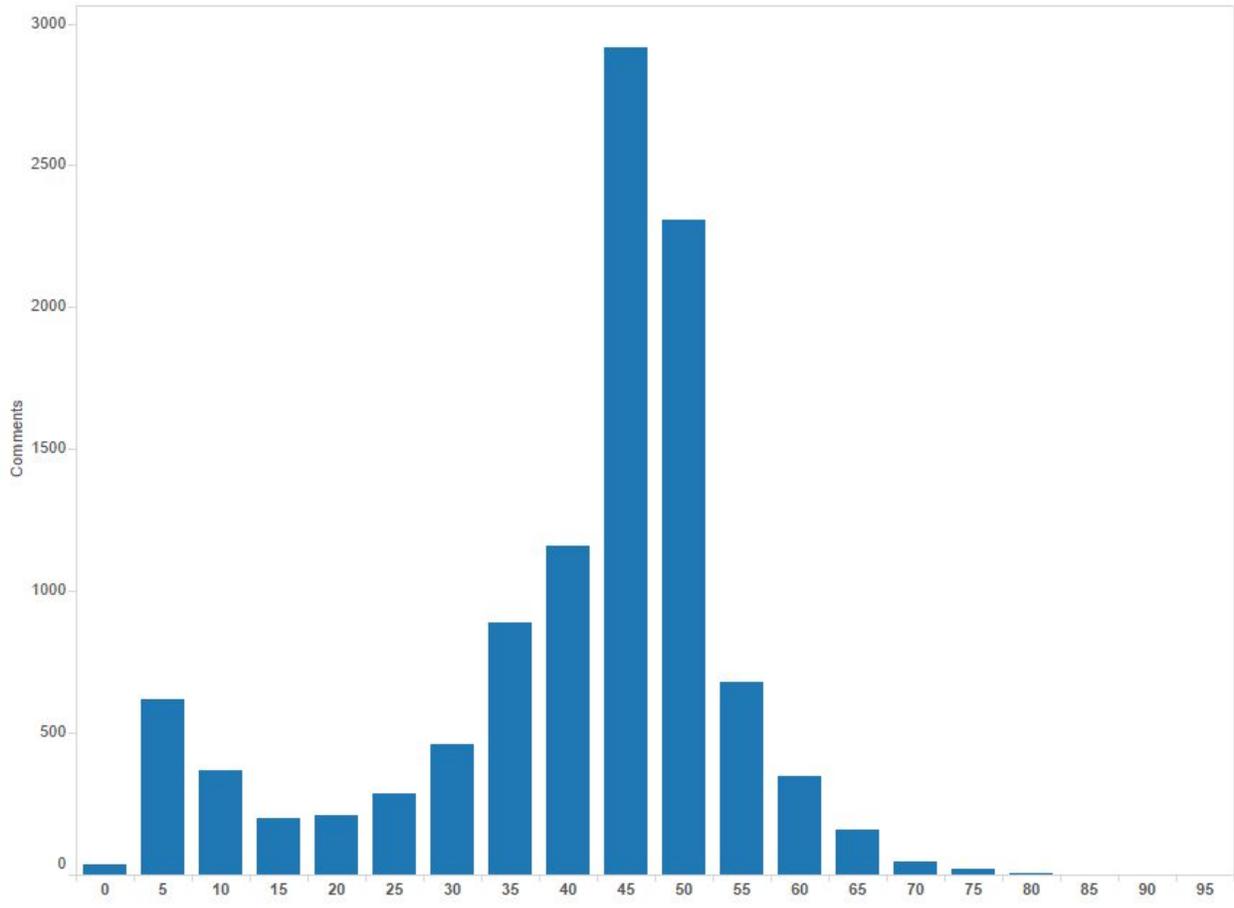


**Table 4: Readability scoring of public comments in response to Docket 2, VA-2016-VHA-0011.**

	<b>FKRE</b>	<b>FKGL</b>	<b>GFOG</b>	<b>SMOG</b>	<b>AGL</b>
<b>Maximum Score</b>	95.9	24.5	28.8	23.8	24.3
<b>Minimum Score</b>	0.1	2.3	3.9	4.9	4.1
<b>Mean</b>	42.8	11.8	14.6	13.6	13.4

<b>Median</b>	47.2	10.6	13.3	12.7	12.5
<b>Mode</b>	52.7	10.2	12.8	12.4	11.5
<b>Standard Deviation</b>	14.5	3.0	3.7	2.6	2.8
<b>Within 1SD of Mean</b>	8668	7434	7390	7349	7723

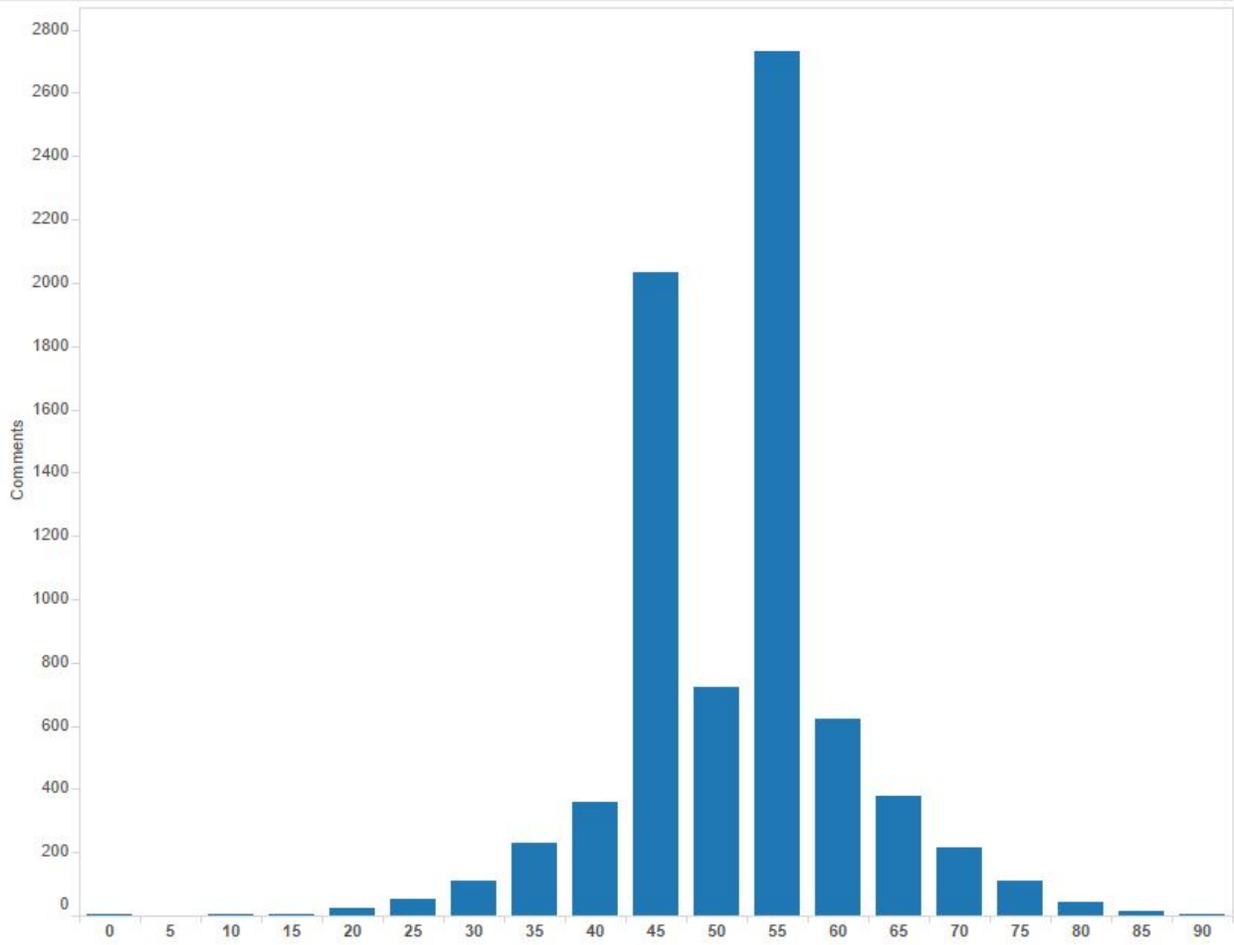
**Figure 3: Comment Frequency Distribution by FKRE Score for Docket 2, VA-2016-VHA-0011.**



**Table 5: Readability scoring of public comments in response to Docket 3, DOS-2014-0003.**

	<b>FKRE</b>	<b>FKGL</b>	<b>GFOG</b>	<b>SMOG</b>	<b>AGL</b>
<b>Maximum Score</b>	94.6	30.9	35.4	23.3	27.9
<b>Minimum Score</b>	0.4	2.2	3.7	4.6	4.2
<b>Mean</b>	54.3	9.79	12.4	11.9	11.3
<b>Median</b>	56.4	9.5	12.2	11.7	10.9
<b>Mode</b>	57.7	11.7	13.8	12.9	12.6
<b>Standard Deviation</b>	9.2	2.5	2.6	1.8	2.0
<b>Within 1SD of Mean</b>	6158	6093	6114	6061	5923

**Figure 4: Comment Frequency Distribution by FKRE Score for Docket 3, DOS-2014-0003.**

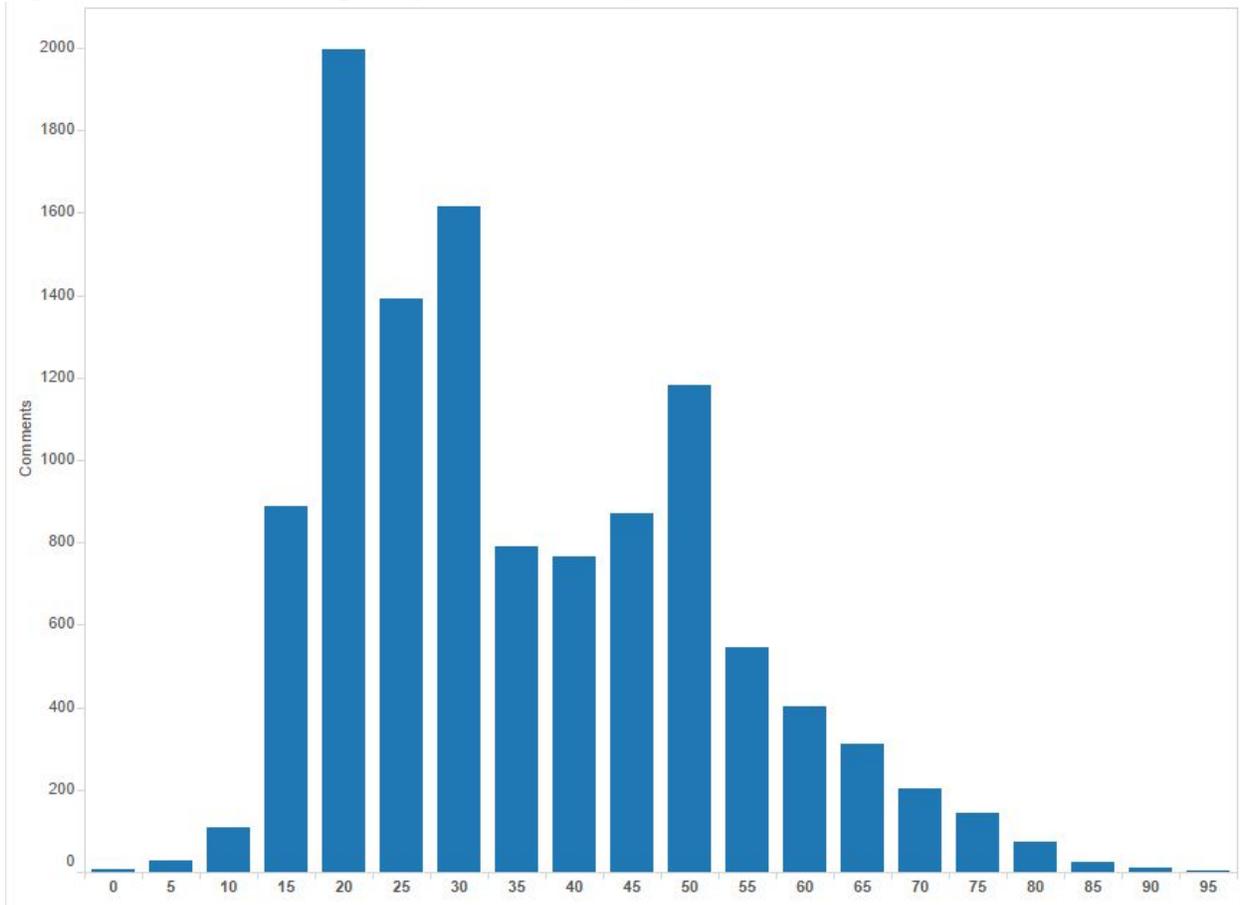


**Table 6: Readability scoring of public comments in response to Docket 4, IRS-2013-0038.**

	<b>FKRE</b>	<b>FKGL</b>	<b>GFOG</b>	<b>SMOG</b>	<b>AGL</b>
<b>Maximum Score</b>	99.8	31.8	34	23.5	26
<b>Minimum Score</b>	0.5	0.9	2.1	3.9	2.9
<b>Mean</b>	38.0	13.0	16.6	14.9	14.5
<b>Median</b>	34.6	13.2	16.8	15	14.7
<b>Mode</b>	34.6	12.3	16.5	14.7	14.7
<b>Standard Deviation</b>	15.9	3.1	3.7	2.7	3.0
<b>Within 1SD of Mean</b>	6686	7937	7842	7911	7850



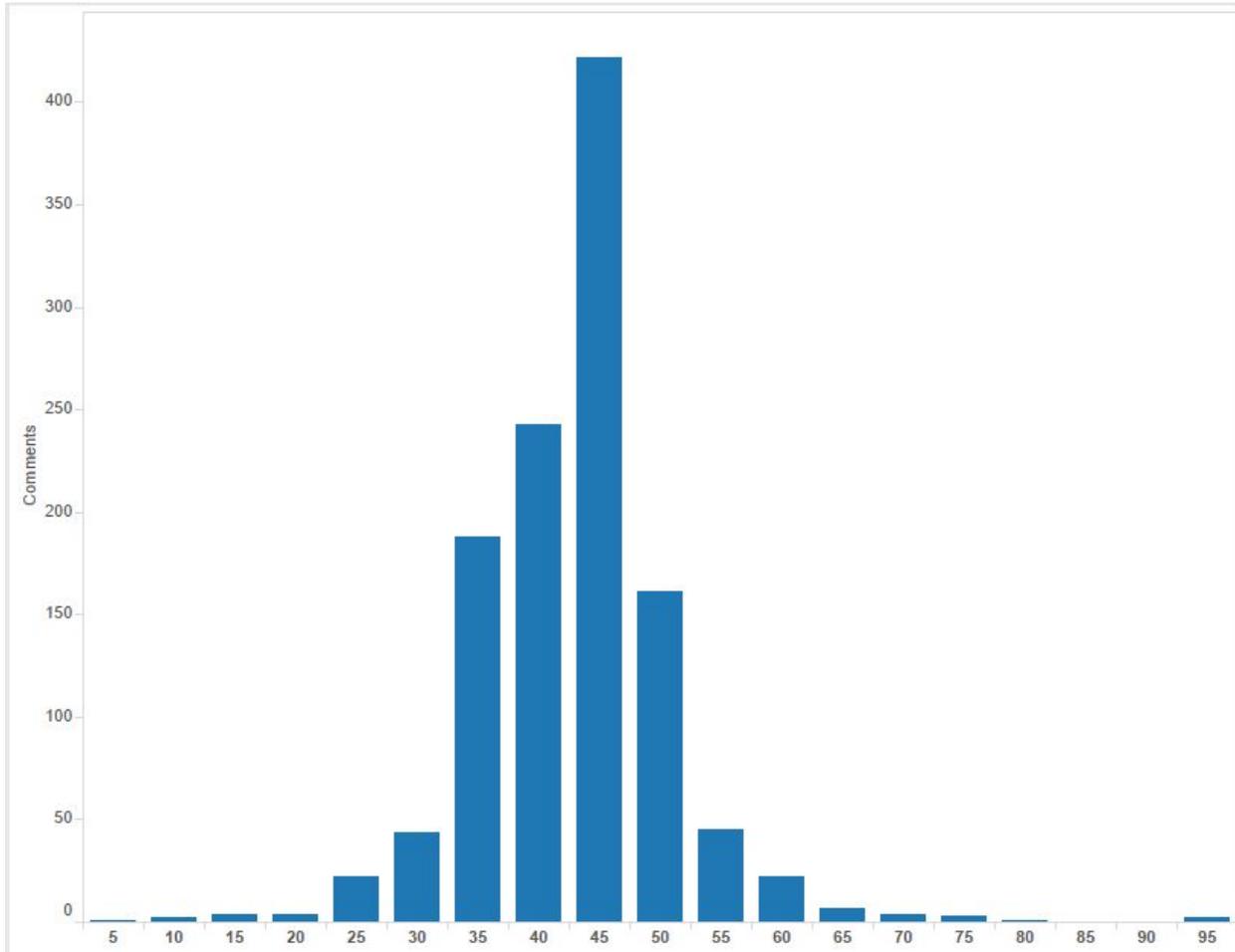
**Figure 5: Comment Frequency Distribution by FKRE Score for Docket 4, IRS-2013-0038.**



**Table 7: Readability scoring of public comments in response to Docket 5, HHS-OS-2016-0014.**

	<b>FKRE</b>	<b>FKGL</b>	<b>GFOG</b>	<b>SMOG</b>	<b>AGL</b>
<b>Maximum Score</b>	95.4	22.4	26.4	21.2	22.1
<b>Minimum Score</b>	11.5	2.1	3.6	4.5	3.6
<b>Mean</b>	45.3	12.1	15.2	13.9	13.8
<b>Median</b>	46	11.7	15.4	14.1	13.8
<b>Mode</b>	46	11.6	15.8	14.5	14
<b>Standard Deviation</b>	7.6	2.0	1.8	1.3	1.5
<b>Within 1SD of Mean</b>	910	763	944	968	918

**Figure 6: Comment Frequency Distribution by FKRE Score for Docket 5, HHS-OS-2016-0014.**



**Table 8: Scoring Variation in Analysis of Docket 5, HHS-OS-2016-0014, Title X Funding**

Scoring Engine	Flesch-Kincaid Reading Ease	Flesch-Kincaid Grade Level	Words	Sentences
<b>Software-Based Tools</b>				
Flesh 2.0	29.35	14.73	9111	407
Microsoft Word	22.2	16.8	6672	227
<b>Online Tools</b>				
Readability-Score.com	39	12	6509	383
Datayze.com	33.95	12.87	6506*	372
Online-Utility.Org	34.54	12.98	6801	372
WebPageFx.com	40.3	12.3	6800	366

\* Note: Datayze.Com Number of words extrapolated from Sentence and Sentence Length.

**Table 9: Flesch-Kincaid Readability of Docket Items**

	<b>FKRE</b>	<b>FKGL</b>
Docket 1: CFPB-2016-0025	27.8	17.0
Docket 2: VA-2016-VHA-0011	38.5	12.4
Docket 3: DOS-2014-0003	51.1	8.6
Docket 4: IRS-2013-0038	30.3	14.2
Docket 5: HHS-OS-2016-0014	39.0	12.0
<b>Mean, All Dockets</b>	<b>37.3</b>	<b>12.8</b>