Opinion Clarity in State and Federal Trial Courts

Adam Feldman*

*To whom correspondence should be addressed.

Abstract

Judges' main output is their writing in the form of judicial opinions. Thousands of opinions are released annually by courts in the United States, with a majority coming from courts of first impression at both state and federal levels. These written opinions explain the law for the parties to a case as well as for similarly situated individuals. Lower court opinions also play substantial roles in cases' appeals. Not all judges do an equal job of clearly articulating analyses of the law and not all cases require the same depth of analysis. This chapter uses automated text analysis software to assess the clarity of judicial opinions from a stratified random sample of state and federal lower courts. After scoring each opinion along a range of metrics for writing clarity, factor analysis was employed to boil down the measures to their salient components. The chapter proceeds by showing how the factors can be employed for quantitative analyses of judicial opinion clarity.

*6677 Colgate Ave., Los Angeles, CA 90048; adfeld@gmail.com

Introduction

Style matters in judicial opinions and has been much discussed in a qualitative fashion by commentators and judges themselves, but judges' writing style has been little studied in a quantitative fashion. One impediment in qualitative accounts is that they often lack clear parameters and may leave the evaluator too much discretion.

Judge Posner (1995, 1429-30), for example, separates judicial writing into the "pure" and "impure" styles where "pure" is the more conventional judicial style of writing, "us[ing] technical legal terms without translation into everyday English" while those writing in the impure style, "pretend that what they are doing when they write a judicial opinion is explaining to a hypothetical audience of laypersons why the case is being decided in the way that it is." This dichotomous understanding of judicial writing provides a starting point for examining judicial style. The distinction between writing types that Posner imposes, however, still leaves much subjective choice to reader to define a writing style's positive attributes and whether an opinion is well written.

Qualitative analyses also provide examples of well-written judicial prose. In this vein, Richard Wydick (1994, 5) cites Cardozo's writings describing characteristics such as: an "economy of words," "no archaic phrases," "no misty abstractions," verbs in the "active voice," and "no ambiguities to leave us wondering who did what to whom." These descriptors all relate to the concept of clear writing and begin to establish specific features that relate to this concept.

Trial judges are tasked with deciding cases and writing opinions. Judges' superiors and colleagues may form views of them based on their written opinions. The power of the courts also rests in others' abilities to construe and implement judges' decisions. That is all to say, judicial writing clarity matters in two directions — it matters to judges and it matters to the laws' recipients. For this reason, multiple individuals are consistently concerned with the clarity of judicial opinion writing.

With increased technological innovation, we are now at a point where we can quantitatively measure aspects legal writing. Although we lack a universal gold standard for "good" judicial writing, certain conventions are both measurable and widely accepted as important. For many reasons, clear writing is generally held as one of several yardsticks by which we can gauge the quality of judicial writing.

Judicial writing clarity makes a difference to those who are affected by the law and to those who implement the law and based on this judges are incentivized to write clearly. From judges' perspectives, they want to be understood and want favorable appraisals of their writings while readers need to make sense of the writings and of the potential consequences that flow from them (Baum 2009, 109).

Judges may particularly care about their colleagues' thoughts and perceptions of their writings (50-52), and if so they will want to tailor their opinions to meet their colleagues' expectations.

They may even be concerned that their future prospects of promotion depend on favorable evaluations of their opinions from other judges and institutional higher-ups (Posner 2004, 1265-73). For state judges facing reelection, this concern may even extend to the general population.

Several commentators have emphasized the fact that judges frequently write for a variety of audiences (Baum 2009; Mikva 1987; Posner 2010). That is, lower court judges might perceive their audiences as different from those of appeals court judges. Lower court judges, for instance, may be worried about how appeals courts will interpret and rule on their decisions. Furthermore, judges opinions in different legal areas may vary based on the intended recipients of the opinions.

On a normative level, interested parties and other individuals that may be affected by certain decisions care about judges' opinions. They need clear direction from judges on the nature of the law, how to conform to the law, and of the consequences (or potential consequences) of their actions.

Since judges do not actually implement their decisions, they must

state the law in sufficiently clear terms so that their decisions can be effected. Judge Patricia Wald (1995, 1372) described that judges write for two reasons: first to "tell others...what to do" and second because, "under a government of laws, ordinary people have a right to expect that the law will apply to all citizens alike." Without clear guidance implementation is impossible. Further, people have no way of acting in accordance with the law if they do not know what the law is.

Unclear legal dispositions may disproportionately affect certain sectors of the population (Yamamoto 1990), and as the audience to which an opinion applies is not always clear at the outset, judicial writing that lacks clarity can have unpredictable downstream consequences.

Judge Posner provides the example of how an opinion dense with citations may be intended to imbue the text with "confidence," but may have the unintended consequence of leaving the law unclear (Posner 1995, 1442-43). In this instance, Judge Posner warns judges against assuming texts steeped in legalese and jargon will be comprehensible by the lay population.

An important caveat though is that laws themselves may be ambiguous to no fault of judges. In Justice Stevens' majority opinion in *Chicago v. Morales* (1999), the Supreme Court declared that people could not be held responsible for following unconstitutionally vague laws which do not give sufficient direction to officials or to the public on what constitutes a crime. The practical problem that arose in *Morales* created by ambiguity in the law is not only a concern for legislatures.

The statutory vagueness present in *Morales* is different from the type of clarity desired in judicial opinions though. Legislatures need to be sure that they clearly establish the population whose behavior is affected by their decrees. Judges on the other hand need to be sure that their writings are comprehensible by a variety of audiences. Similar to legislatures, however, vague judicial writing can lead to constitutional due process concerns.

This chapter looks at the relative clarity of legal writing by quantitatively analyzing a stratified random sample of judicial opinions from state and federal courts of first impression. The analyses in this chapter also build on prior comparative analyses of state and federal courts. Eisenberg et al. (1995) for example found that awards in federal jury trials are greater than those in state courts and that state court cases proceed more slowly than federal court cases. Galanter (2004) showed that trials are declining as a method of dispute resolution with a similar frequency in federal and state courts.

The sample used in this chapter includes a combination of over 4,000 state and federal court opinions. Text analyses were performed on these data to quantify opinion clarity. After focusing on measurement strategies, this chapter implements various metrics to compare the opinions from state and federal courts. This exploratory analysis provides a basic comparison of opinions in both forums so that we can begin to grasp distinctions in opinion clarity. The chapter concludes by discussing how this work can be applied in future studies interested in quantitatively comparing judicial opinions and other types of legal writing.

Prior Work

Although previous works have touched on the topic of legal writing clarity, this is the first time to the author's knowledge that trial court opinions are compared and analyzed along these lines.

Much of the recent quantitative scholarship of judicial writing quality looks at Supreme Court briefs and opinions. Other works expand their loci beyond the United States Supreme Court to forums like state supreme courts (Goelzhauser and Cann 2014). Some works in this domain examine the importance of clearly written Supreme Court opinions and how Supreme Court justices vary their writing clarity depending on their intended audience(s) (Black et al. 2016b; Owens and Wedeking 2011).

Analyses of judicial writings in the Supreme Court have taken several different routes. Carlson, Livermore, and Rockmore (2015) looked for differences in Supreme Court writing styles over time and found more within justice variation in writing style as the justices took on a greater number of clerks over time. Corley and Wedeking (2014) used LIWC's dictionary-based software to uncover that as the Court uses language with greater certainty, lower courts are more likely to treat the Court's decisions favorably.

Several works also examine the public's responses to Supreme Court opinions depending on opinion clarity. Hansford and Coe (2014) use a survey design to obtain these findings. Black et al. (2016) use a similar factor analysis to the one used in this paper. That paper uses the Korpus package in R to generate readability scores which compose the readability index variable created through factor analysis.

Feldman (2016) found that Supreme Court briefs with higher quality writing were more likely to succeed and to affect opinion content. Spencer and Feldman (2018) recently found that according to a similar measure to the one used in this paper also created through factor analysis of readability measures that motions for summary judgment with clearer writing were more likely to succeed in both state and federal trial courts. Prior to this, Long and Christensen (2011) examined whether the readability of Supreme Court briefs affected litigants' likelihood of success and did not find a meaningful relationship between more readable briefs and successful outcomes.

This chapter engages in similar analyses but with different forums of focus and with different expectations about possible outcomes. The following section lays out and describes the data, methodology, and basic hypotheses.

Methodology & Analysis

The data for these analysis consists of a sample of over 4,000 federal and state court opinions. All opinions were obtained through Westlaw. A

random sampling technique was used within product-liability decisions to obtain opinions and to control for issue content. The dates of the opinions range from 2000 through 2015. The search parameters were then set to first include only state trial court cases and then only federal district court cases.

Opinions were saved as plain text files. The Quanteda package in R Markdown was then used to generate readability metrics for each opinion.

Clarity Scores

Quanteda in R provides multiple text statistics metrics including several different readability scores. The readability scores include: Flesh-Kincaid Grade Level, Gunning Fog Score, Coleman Liau Index, Smog Index, Automated Readability Index, and Spache Scores. These algorithms differ in how they use and weight certain linguistic factors such as word and sentence counts and lengths and so they each answer the question of what makes a text clearer or more readable from slightly different angles.

Each of the underlying readability measures was designed for a different purpose and application. Each of these metrics has its own units of measurement although most track grade level fairly closely with higher grade levels equating to more complicated and less clear texts. All of these underlying scores move in the same direction.

Flesch-Kincaid scores were developed for use by the navy in the mid-1970's (Kincaid et al. 1975). They have since been used in applications ranging from evaluating the ease of reading naval training manuals (Kincaid, Aagard, and O'Hara 1980) to hospitals' patient information leaflets (Williamson and Martin 2010). The Automated Readability Index (ARI) was also created to gauge the readability of training manuals (in the case of ARI these were Air Force manuals) (Senter and Smith 1967) and has been employed within a variety of industries and even to measure the readability of online product reviews (Korfiatis, Garcia-Bariocanal, and Sanchez-Alonso 2012).

The Simple Measure of Gobbledygook (SMOG) Index, designed for applications in the field of psychology, similarly estimates the appropriate age one should attain to read a given passage of writing (Mc Laughlin 1969). It has been used in a diverse set of applications including measuring the readability of job analysis questionnaires (Ash and Edgell 1975).

The Gunning Fog Index was created in order to estimate the appropriate grade level one should attain to read given writings. Its applications have ranged from measuring the readability of newspaper excepts to business materials (Gunning 1969). The Gunning Fog Index categorizes words with three or more syllables as complex.

The Coleman Liau Index was one of the first readability measures designed for use on a computer and focuses on characters in a word rather than syllables. It also is based on reading grade level and was designed to help the United States Department of Education tailor textbooks to appropriate reading levels for the intended students (Coleman and Liau 1975).

Lastly, the Spache Index was created to gauge the age appropriateness of school reading materials (Spache 1953). Spache Scores each have an associated set of words that are applied in the algorithms. These include over 900 common words such as , "about," "child," and "made." Difficult words are coded as those that are not covered by the list of common words.

The algorithms for each of the readability metrics are as follows:

1. Gunning Fog Index: .4((total words/total sentences) + 100(complex words/words))

2. Coleman Liau Index: .0588(average number of letters per 1,000 words) - .296(average number of sentences per 100 words) - 15.8

3. Smog Index: 1.0430($\sqrt{polysyllables}$) * 30/total sentences) + 3.1291

4. Automated Readability Index (ARI): 4.71(characters/words) + .5(words/sentences) - 21.43

5. Flesch Kincaid Grade Level: .39(total words / total sentences) + 11.8 (total syllables / total words) - 15.59

6. Spache Score (using a sample of text 100-150 words in length): .141(total words/total sentences) + .086(total words/difficult words) + .839

This chapter uses these six readability measures to create an index variable for opinion clarity in the Stata statistical package using factor analysis. Such a factor is described as, "[a] combination of all relevant variables in a particular measurement domain" (Cudeck 2000, 268). These data allow for various tests and comparisons between the state and federal samples as they provide some of the first quantitative distinctions between federal and state trial court judicial writing clarity.

Factor Analysis

Factor analysis is used to parsimoniously represent a set of correlated variables or scores (Brown 2014). The process is defined as, "a set of statistical procedures designed to determine the number of distinct constructs needed to account for the pattern of correlations among a set of measures" (Fabrigar and Wegener 2011). Factor analysis then provides information about the factors that underpin a set of correlated measures. The estimates of influence of the various measures calculated through factor analysis are known as factor loadings.

Factor analysis may lead to the generation of multiple factors if the scores cannot be properly captured with just one factor. This generally is the case when individual variables are particularly unique from one another and so the correlation between them is low.

While factor analysis may yield multiple factors if the underlying values present more than one dimension of results, factor analysis in this instance only generated a single factor.

The six readability variables described above are used in this chapter to derive a common factor through factor analysis. The following section provides a check on the validity of this procedure by applying this method to several non-judicial and publicly accessible texts.

External Validity Check

To first assess the validity of these readability measures and their combination through factor analysis, this procedure was applied to several non-legal texts. Three sets of texts were used for this assessment. The first are three childrens' stories. The second set is comprised of three of the Federalist Papers. The third set includes introductory chapters from machine learning and organic chemistry textbooks.

If one was ordering these works in terms of likely relative writing clarity, one might expect childrens' stories to be the most clear or easiest to read, followed by the textbooks which ostensibly should be easy for laypeople to comprehend, and the most difficult should be the Federalist Papers.

Isolating the text from these eight documents and applying factor analysis led to the creation of six factors. The first factor was the only significant one though as it accounted for over 96% of the variance between the underlying variables. The percentage of variance varies depending on the underlying sample of texts.

The factor loadings for the six variables show that they are all strongly associated with the latent variable. The loadings for five of the variables was over .99 and the sixth measure, Coleman Liau, had a loading of .95. These loadings also vary depending on how they perform measuring the underlying texts.

The underlying variables' uniqueness measures provides more evidence that the variables measure a similar characteristic of writing clarity. These variables' uniqueness scores varied from .001 to .06.

The variables then each get a scoring coefficient that relates the amount of strength the underlying variable has on the latent measure. The scoring coefficients for a factor add up to one. Each of the scoring coefficients for the six underlying variables was within the range of .164 to .172.

This factor analysis was run on the eight texts described above to help gauge whether this method assesses the relative writing clarity of the various pieces in accordance with their hypothesized writing clarity levels.

The clarity scores derived through factor analysis ran from -1.4 to 1.06 with lower scores equating to clearer writing.

The clarity scores for these six documents lined up in the expected manner. The three documents that scored easiest to read were the children stories. The scores for these stories were -1.4, -1.2, and -.76. Next were the textbook chapters. Their scores were .24 and .29. Finally the scores for the Federalist Papers were the highest at .77, 1.04, and 1.06.

The relative reading ease of these texts is is also apparent when examining them qualitatively. The childrens' stories tend to have short sentences with simple vocabulary. One example from the story "The Three Little Pigs," reads,

> "The first little pig was very lazy. He didn't want to work at all and he built his house out of straw. The second little pig worked a little bit harder but he was somewhat lazy too and he built his house out of sticks."

The sentences in the textbooks are also relatively short and to the point. The machine learning text opens, for example,

> "The subject of this book is automated learning, or, as we will more often call it, Machine Learning (ML). That is, we wish to program computers so that they can 'learn' from input available to them. Roughly speaking, learning is the process of converting experience into expertise or knowledge."

There is no excessive jargon in these introductory chapters and they are written for a non-technical audience.

The sections from the Federalist Papers are written in the most complex manner of the three sets of texts and are the least clear as they are the most difficult to comprehend. Here is a snippet of text from Federalist Paper 1 as an example:

> "And yet, however just these sentiments will be allowed to be, we have already sufficient indications that it will happen in this as in all former cases of great national discussion. A torrent of angry and malignant passions will be let loose. To judge from the conduct of the opposite parties, we shall be led to conclude that they will mutually hope to evince the justness of their opinions, and to increase the number of their converts by the loudness of their declamations and the bitterness of their invectives."

The text has multiple clauses, complicated language, and is the most difficulty to follow of the three sets of writings.

Clarity of Judicial Opinions

What do legal writings high and low on the clarity index look like? An example of a sentence from an opinion in the dataset that scored high according to this index (i.e. lower in clarity) is:

> "The aforementioned Defendants moved the Court to enter an order precluding Plaintiffs counsel from referring to, mentioning, engaging in commentary and/or otherwise eliciting testimony, whether direct or indirect, concerning the amount of profits or royalties derived from any of the operations that occurred on the Defendants' property and the amount of consideration paid by the Defendants for the properties at issue."[1]

1. *In re: Flood Litigation*, 2006 WL 6460516 (W.Va. Cir. 2006).

On average, sentences from opinions with higher scores are lengthier with more clauses and longer words, while sentences from opinions with lower scores are more direct, less loquacious, and more to the point.

The following excerpt provides several sentences from one of the clearest written opinions based on its clarity score:

> "1. The accident, which forms the basis of Plaintiff's Amended Complaint, occurred on December 12, 1995, at the construction site of Solar Sources Inc.'s (hereinafter 'Solar') coal preparation plant in Cannelburg, Indiana. 2. At the time of the accident, the construction of the coal preparation plant was approximately halfway completed. (Deposition of Steven L. Vaughn taken on February 26, 1999, p. 191, hereinafter 'Vaughn Dep.'). 3. Plaintiff was employed as a pipe-fitter by Trimble Engineers and Constructors."[2]

This passage lays out the case's fact pattern with discrete statements. The sentences are easy to follow as they flow from one to the next and this is in sharp contrast to the sentence from the earlier opinion that scored higher according to the clarity measure.

To derive clarity scores, a similar factor analysis was performed on the over 4,000 opinions in the dataset based on the six underlying variables for writing clarity. The computation of the factor variable was somewhat different for these opinions compared with the computation for the texts in the validity check.

Although the factor analysis only located one significant factor, the single factor in this instance accounted for 74% of the variance between the underlying variables.

The factor loadings are quite different in this instance from those in the validation set. Two of the underlying variables, Coleman Liau and

2. *Vaughn v. Daniels Company*, 2001 WL 36039333 (Ind. Cir. 2001).

Gunning Fog have negative factor loadings of -.86 and -.65 respectively. The other factors have positive loadings: .75 for SMOG, .93 for ARI, and .96 for both Flesch Kincaid and Spache scores.

These loadings lead the latent variable to take on more disparate scoring coefficients for the six measures. The scoring coefficients in this instance are -.19 for Coleman Liau, -.15 for Gunning Fog, .17 for SMOG, .21 for ARI, .22 for Flesch Kincaid, and .22 for Spache.

The values for the clarity scores were derived based on how the factor analysis balanced the underlying variables when creating a common factor. As the underlying scores all increase when the reading level of the sample text increases, the clarity scores increase as a text's readability decreases.

To get a lay of the land, Table 1 below provides the descriptive statistics for the clarity scores as they are dispersed about the dataset.

Table 1. *Summary Statistics for Clarity Scores*

Federal		State		Overall	
N	2,100	N	1,946	N	4,046
25th percentile	0.49	25th percentile	-1.03	25th percentile	-0.92
75th percentile	0.95	75th percentile	-0.81	75th percentile	0.69
99th percentile	2.81	99th percentile	-.29	99th percentile	2.45
Mean	0.81	Mean	-0.88	Mean	0.00
Median	0.68	Median	-0.93	Median	0.21
Sd. Dev.	0.69	St. Dev.	0.30	St. Dev.	1.0

The range for all clarity scores was from -1.40 to 15.0 with all but three scores under 5.79 and with lower scores equating to more readable texts. From these data, there are readily apparent differences between the statistics for state and federal opinions. The state scores, for instance tend to fall below those for the mixed federal and state (overall) values while

the federal scores are generally above this threshold.[3] The spread of the data is slightly greater in the federal sample as its standard deviation is .69 while this measure for the state sample is .30.

One explanation for a portion of the difference is that federal product-liability cases are predicated on underlying state law claims where cases have been transferred to federal court based on diversity or federal question jurisdiction. This added case complexity can explain some of the difference between federal and state court clarity.[4]

To help better visualize the set of clarity scores, the two histograms below show the distribution of values for the two sets.

The clarity scores for the state opinions are clearly to the left of those for federal courts. This suggests that federal court opinions in the sample tend to be less clear than state court opinions. The federal court's scores also extend much farther to the right in tail.

Additional analyses confirm that the differences between the state and federal distributions are significant. A two-sample Kolmogorov-Smirnov test for the equality of distribution showed that state court opinions' clarity scores tended to be smaller than those for federal courts at less than a 0.01 p-level with the largest difference between the distribution functions as 0.97.

3. As a robustness check, a two-sample Kolmogorov-Smirnov test was run and verified that state opinions' words per sentence were smaller than those for federal court opinions and this result was significant at p-level of less than 0.01.

4. Samples of several paragraphs that did not relate to transfer jurisprudence from 30 federal opinions and from 30 state opinions were used to test this hypothesis. While the median value for state opinions was still lower the difference in median values is considerably less than it is for the entire sample.

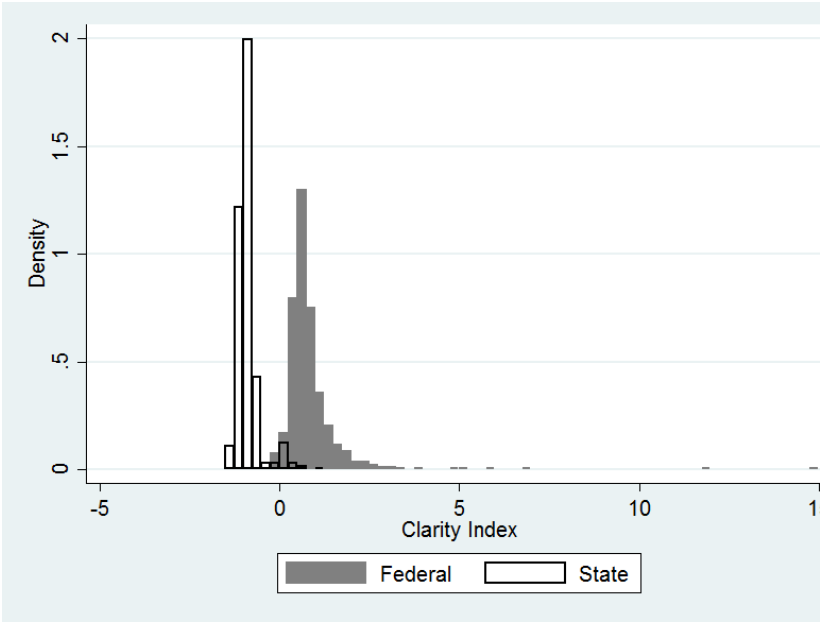

Figure 1. *Federal and State Clarity Score Histograms*

Word Counts

Federal and state court opinions in the sample differ in their length. Much of this difference may relate to the added complexity in cases transferred from state to federal court based on diversity jurisdiction.

To further verify this graphical finding on opinion length, a two-sample Kolmogorov-Smirnov test for whether state opinion lengths were shorter than those for federal courts was run as a robustness check was significant at a p-value of less than .01.

The median opinion length for the entire set of federal and state court opinions is 1,532 words. The median length for state opinions, however, is 790 words while the median length for federal court opinions is 2119.5 words.

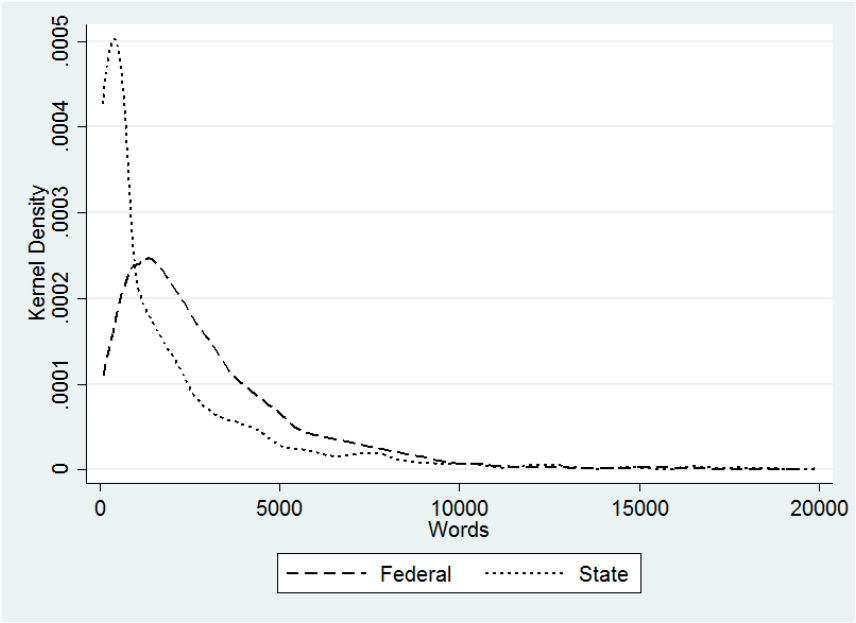

Figure 2. *Kernel Density Plot of Federal and State Court Opinion Word Length*

Note: 25 opinions over 20,000 words removed from plot.

Overly lengthy judicial opinions have been a point of concern of jurists and commentators for decades (McComb 1949). Further highlighting this point, Gerald Lebovitz (2004, 64) writes of the virtues of shorter opinions, "Brief opinions hold the reader's attention, allow readers to move on to other things, and distill the opinion's essence." Thus, longer opinions may stymie readers' attempts to make sense of legal text and prevent readers' from recognizing the salient points in opinions even if judges have no choice but to write lengthy opinions given the facts and law involved in a cases.

The variation in length between federal and state court opinions may not relate to judges' volitional choices. Many of the state court opinions appear shorter because they were summarily decided in a few para-

graphs. Such decisions do not involve the complex legal analyses that are present in more extensive opinions. These aspects are in additional to the more detailed analyses required of federal court judges examining whether it is appropriate for them to take cases from state courts based on removal.

Conclusion

This chapter presents one of the first endeavors to apply such metrics of writing clarity to federal and state court trial court opinions. It uses exploratory analysis to uncover aspects of judicial opinions that are related to writing clarity.

Much of the difference in writing clarity between federal and state court judges' opinions may have little to do with the judges themselves and a lot to do with added complexity of federal cases, especially related to jurisdictional concerns. This is especially the case if the transfer analysis in federal opinions helps drive the lower clarity of federal opinions or if the summary decisions in state courts help drive the greater clarity in state court opinions.

Looking beyond the rationale for the difference in clarity for a moment, multiple audiences still need to be able to comprehend both federal and state court judges' writings. As a judge's key output along with decisions, written opinions represent judges to the outside world. Lawyers, academics, and other judges form views about judges based on their written opinions. The public relies on these opinions to guide their actions in accordance with the law. Judges seeking promotion are also incentivized to write clear opinions in the event that their work is assessed in the course of the promotion process. Other reasons could also be offered for judges to focus on their writing clarity such as a duty to the parties.

Judicial opinions affect parties to the case and may also affect a broader population. There are vast implications if opinions at the trial court level, where the bulk of litigation occurs, are not clear. Ultimately bodies outside of the courts are responsible for implementing courts' decisions

(Rosenberg 2008). Since instructions from the courts often direct enforcement bodies' actions, opinions that lack clear writing may impede the functionality of such a system.

The results from this chapter suggest that federal court opinions, at least in the area of products liability, are less clear according to readability standards than state court opinions, although the relative clarity of the opinions in these two samples is not necessarily an inherent reflection on the judges' writing quality.

Although a far less efficient approach in a large sample context, qualitative assessments of opinion clarity can further support such quantitative findings and can help uncover distinctions between texts that are not immediately apparent with quantitative analyses. While the tools for text analysis continue to improve and evolve, there will surely be even more accurate ways to gauge the clarity of legal writing in the near future. We will also likely see tools specifically created for measuring aspects of legal writing rather than general writers' aids. With off-the-shelf approaches like the one in this chapter, quantitatively measuring the clarity of writing is possible and the findings help us assess differences in state and federal court opinions in a manner that would be inaccessible with qualitative analyses alone.

References

Ash, Ronald A, and Steven L Edgell. 1975. "A Note on the Readability on the Position Analysis Questionnaire (PAQ)." *Journal of Applied Psychology* 60 (6): 765.

Baum, Lawrence. 2009. *Judges and Their Audiences: A Perspective on Judicial Behavior*. Princeton University Press.

Black, Ryan C, Ryan J Owens, Justin Wedeking, and Patrick C Wohlfarth. 2016a. "The Influence of Public Sentiment on Supreme Court Opinion Clarity." *Law & Society Review* 50 (3): 703–732.

Black, Ryan C, Ryan J Owens, Justin Wedeking, and Patrick C Wohlfarth. 2016b. *US Supreme Court Opinions and Their Audiences.* Cambridge University Press.

Brown, Timothy A. 2014. *Confirmatory Factor Analysis for Applied Research.* Guilford Publications.

Carlson, Keith, Michael A Livermore, and Daniel Rockmore. 2015. "A Quantitative Analysis of Writing Style on the US Supreme Court."

Coleman, Meri, and Ta Lin Liau. 1975. "A Computer Readability Formula Designed for Machine Scoring." *Journal of Applied Psychology* 60 (2): 283.

Corley, Pamela C, and Justin Wedeking. 2014. "The (Dis)advantage of Certainty: The Importance of Certainty in Language." *Law & Society Review* 48 (1): 35–62.

Cudeck, Robert. 2000. "Exploratory Factor Analysis." *Handbook of Applied Multivariate Statistics and Mathematical Modeling* 265:296.

Eisenberg, Theodore, John Goerdt, Brian Ostrom, and David Rottman. 1995. "Litigation Outcomes in State and Federal Courts: A Statistical Portrait." *Seattle UL Rev.* 19:433.

Fabrigar, Leandre R, and Duane T Wegener. 2011. *Exploratory Factor Analysis.* Oxford University Press.

Feldman, Adam. 2016. "Counting on Quality: The Effects of Merits Brief Quality on Supreme Court Decisions." *Denv. L. Rev.* 94:43.

Galanter, Marc. 2004. "The vanishing trial: An Examination of Trials and Related Matters in Federal and State Courts." *Journal of Empirical Legal Studies* 1 (3): 459–570.

Goelzhauser, Greg, and Damon M Cann. 2014. "Judicial Independence and Opinion Clarity on State Supreme Courts." *State Politics & Policy Quarterly* 14 (2): 123–141.

Gunning, Robert. 1969. "The Fog Index After Twenty Years." *Journal of Business Communication* 6 (2): 3–13.

Hansford, Thomas G, and Chelsea Coe. 2014. "Linguistic Complexity and Public Acceptance of Supreme Court Decisions." *Washington, DC: AMAPSA*.

Kincaid, J Peter, James A Aagard, and John W O'Hara. 1980. *Development and Test of a Computer Readability Editing System (CRES).* Technical report. Training Analysis and Evaluation Group (Navy) Orlando FL.

Kincaid, J Peter, Robert P Fishburne Jr, Richard L Rogers, and Brad S Chissom. 1975. *Derivation of New Readability Formulas for Navy Enlisted Personnel.* Technical report. Naval Technical Training Command Millington TN Research Branch.

Korfiatis, Nikolaos, Elena Garcia-Bariocanal, and Salvador Sanchez-Alonso. 2012. "Evaluating Content Quality and Helpfulness of Online Product Reviews: The Interplay of Review Helpfulness vs. Review Content." *Electronic Commerce Research and Applications* 11 (3): 205–217.

Lebovits, Gerald. 2004. "Short Judicial Opinions: The Weight of Authority." *NY St. B. Assn. J.* 76:64.

Long, Lance N, and William F Christensen. 2011. "Does the Readability of Your Brief Affect Your Chance of Winning an Appeal." *J. App. Prac. & Process* 12:145.

Mc Laughlin, G Harry. 1969. "SMOG Grading-a New Readability Formula." *Journal of reading* 12 (8): 639–646.

McComb, Marshall F. 1949. "A Mandate from the Bar: Shorter and More Lucid Opinions." *American Bar Association Journal:* 382–384.

Mikva, Abner J. 1987. "For Whom Judges Write." *S. Cal. L. Rev.* 61:1357.

Owens, Ryan J, and Justin P Wedeking. 2011. "Justices and Legal Clarity: Analyzing the Complexity of US Supreme Court Opinions." *Law & Society Review* 45 (4): 1027–1061.

Posner, Richard A. 1995. "Judges' Writing Styles (And Do They Matter?)" *The University of Chicago Law Review* 62 (4): 1421–1449.

———. 2004. "Judicial Behavior and Performance an Economic Approach." *Fla. St. UL Rev.* 32:1259.

———. 2010. *How Judges Think.* Harvard University Press.

Rosenberg, Gerald N. 2008. *The Hollow Hope: Can Courts Bring About Social Change?* University of Chicago Press.

Senter, RJ, and Edgar A Smith. 1967. *Automated Readability Index.* Technical report. CINCINNATI UNIV OH.

Spache, George. 1953. "A New Readability Formula for Primary-Grade Reading Materials." *The Elementary School Journal* 53 (7): 410–413.

Spencer, Shaun B, and Adam Feldman. 2018. "The Empirical Relationship between Brief Clarity and Summary Judgment Decisions." *Legal Writing: Journal of the Legal Writing Institute* 22.

Wald, Patricia M. 1995. "The Rhetoric of Results and the Results of Rhetoric: Judicial writings." *The University of Chicago Law Review* 62 (4): 1371–1419.

Williamson, James Matthew Lloyd, and AG Martin. 2010. "Analysis of Patient Information Leaflets Provided by a District General Hospital by the Flesch and Flesch–Kincaid Method." *International journal of clinical practice* 64 (13): 1824–1831.

Wydick, Richard C. 1994. *Plain English for Lawyers.* Vol. 4. Carolina Academic Press.

Yamamoto, Eric K. 1990. "Efficiency's Threat to the Value of Accessible Courts for Minorities." *Harv. CR-CLL Rev.* 25:341.