



LegalWriting  
institute

## **Monograph Series**

### **Volume 13—Examining Legal Writing Empirically**

This article was originally published with the following citation:

Shaun Spencer, *Using Empirical Methods to Study Legal Writing*, 20 LEGAL WRITING 141 (2015).

Reprinted with permission.

# USING EMPIRICAL METHODS TO STUDY LEGAL WRITING

Shaun B. Spencer\*

The goal of this Article is to promote an emerging field of legal writing scholarship: the empirical study of legal writing. Reading this Article should not make you an expert in qualitative or quantitative research methods. Instead, it should highlight the possibilities for future empirical legal writing scholarship and offer enough of an introduction to inspire new empirical researchers.<sup>1</sup> Even if you do not plan to conduct your own empirical research, this Article should make you a more informed consumer of empirical scholarship.

This Article does not suggest that empirical research is easy. However, with careful attention to methodology, empirical research methods can yield potentially valuable findings. Given the talent and energy among legal writing faculty, we are well positioned to study what lawyers write, and the lawyers who read and write it.

Part I introduces empirical research methods. This part first locates empirical methods within the existing fields of legal writing scholarship and then offers an overview of the three empirical research strategies: qualitative, quantitative, and mixed methods research. Parts II, III, and IV examine how researchers can apply quantitative, qualitative, and mixed methods research strategies to study legal writing. Each part describes key features of each

---

\* Assistant Professor and Director of Legal Skills, University of Massachusetts School of Law–Dartmouth. The Author received many helpful suggestions on this topic during a presentation at the 16th Biennial Conference of the Legal Writing Institute on June 29, 2014. The Author is grateful for invaluable revisions by the Editorial Board of the *Legal Writing Journal*, and for the support of the UMass Law Summer Research Grant Program.

1. For earlier work on empirical research and legal writing, see Sarah J. Morath, *It's Not All Statistics: Demystifying Empirical Research*, 27 *SECOND DRAFT*, Summer 2013, at 22–24, available at, [http://lwionline.org/uploads/FileUpload/LWISecundDraft\\_vol27no1b.pdf](http://lwionline.org/uploads/FileUpload/LWISecundDraft_vol27no1b.pdf); Robin A. Boyle & Joanne Ingham, *Suggestions on How to Conduct Empirical Research: A Behind-the-Scenes View*, 15 *PERSPS*. 176 (2007) (focusing on quantitative methods).

research method and illustrates those features by describing prior empirical studies of legal writing. Part V offers practical advice to new empirical researchers about potential questions for empirical research, data collection issues, the role of institutional review boards and methodologists, and writing up a study. Finally, the Appendix is a bibliography that presents general references on empirical research methods and lists existing empirical studies of legal writing.

### I. INTRODUCTION TO EMPIRICAL RESEARCH

Legal writing scholarship has existed in some form for nearly a century but has exploded since the 1980s with the support of the newly formed Legal Writing Institute, biennial conferences, and several legal writing journals and newsletters.<sup>2</sup> Terrill Pollman and Linda Edwards described four common legal writing topics in contemporary legal writing scholarship: “those related to (1) the substance or doctrine legal writing professors teach; (2) the theories underlying that substance; (3) the pedagogy used to teach that substance; and (4) the institutional choices that affect that teaching.”<sup>3</sup> Empirical research offers another valuable branch of study in this dynamic and evolving discipline. In fact, in their recently published book, Catherine Cameron and Lance Long explain how empirical studies in other disciplines as well as in the legal writing field can inform the work of legal writers.<sup>4</sup>

People often equate the phrase “empirical research” with statistical analysis. But that refers to only one branch of empirical research—quantitative research. In contrast, qualitative research often involves an in-depth exploration of individual experiences and understandings. If quantitative observations involve “hard” data like statistical analysis of variables, then qualitative observations involve “soft” data like field observations and interviews.<sup>5</sup>

---

2. Linda L. Berger et al., *The Past, Presence and Future of Legal Writing Scholarship: Rhetoric, Voice, and Community*, 16 *LEGAL WRITING* 521, 525–33 (2010); George D. Gopen & Kary D. Smout, *Legal Writing: A Bibliography*, 1 *LEGAL WRITING* 93, 93 (1991).

3. Terrill Pollman & Linda H. Edwards, *Scholarship by Legal Writing Professors: New Voices in the Legal Academy*, 11 *LEGAL WRITING* 3, 19 (2005).

4. CATHERINE J. CAMERON & LANCE N. LONG, *THE SCIENCE BEHIND THE ART OF LEGAL WRITING* 6 (2015).

5. JOHN W. CRESWELL, *RESEARCH DESIGN: QUALITATIVE, QUANTITATIVE, AND MIXED METHODS APPROACHES* 4 (4th ed. 2014) [hereinafter CRESWELL, *RESEARCH DESIGN*].

Quantitative research generally focuses on the big picture by studying a sample drawn randomly from a population.<sup>6</sup> Researchers collect quantitative data from the sample and examine the relationships among variables.<sup>7</sup> They quantify the variables through counts (e.g., number of words per sentence, percentage of respondents who prefer one brief over another), categories (e.g., federal or state court, plaintiff or defendant), or ratings (e.g., agreement or disagreement rated on scale of 1 to 5).<sup>8</sup> Quantitative research is predominantly deductive.<sup>9</sup> Researchers begin with a hypothesis, collect variables to test the hypothesis, and draw conclusions about what the relationships between variables mean.<sup>10</sup> Studying legal writing with quantitative methods recognizes that patterns emerge when large groups of people engaged in similar interactions.<sup>11</sup>

Qualitative research, in contrast, seeks a deep, contextual understanding grounded in the participants' context.<sup>12</sup> Researchers gather data through personal interactions or observations, rather than through objective instruments.<sup>13</sup> Common qualitative data collection techniques include observation, interviews, and document review.<sup>14</sup> Qualitative research is primarily inductive.<sup>15</sup> Researchers begin by conducting a close study of the participants in context, and gradually develop themes and theories based on the participants' experiences.<sup>16</sup> Researchers cannot understand what it means to "do" legal writing without understanding the

---

6. *Id.* at 147–49.

7. *Id.* at 151–53.

8. *See id.* at 49–50.

9. *Id.* at 55.

10. *Id.* at 131–32.

11. *See id.* at 4; JOHN W. CRESWELL, EDUCATIONAL RESEARCH: PLANNING, CONDUCTING, AND EVALUATING QUANTITATIVE AND QUALITATIVE RESEARCH 13–16 (4th ed. 2012) [hereinafter CRESWELL, EDUCATIONAL RESEARCH]; GARY SMITH, ESSENTIAL STATISTICS, REGRESSION, AND ECONOMETRICS 141–54 (2012) (discussing samples and populations); R. MURRAY THOMAS, BLENDING QUALITATIVE AND QUANTITATIVE RESEARCH METHODS IN THESES AND DISSERTATIONS 1–6 (2003).

12. CRESWELL, RESEARCH DESIGN, *supra* note 5, at 4, 185–86; MICHAEL QUINN PATTON, QUALITATIVE RESEARCH & EVALUATION METHODS 3–35 (4th ed. 2015).

13. CRESWELL, RESEARCH DESIGN, *supra* note 5, at 4, 185–86; PATTON, *supra* note 12, at 14.

14. CRESWELL, RESEARCH DESIGN, *supra* note 5, at 190–93; PATTON, *supra* note 12, at 14.

15. PATTON, *supra* note 12, at 64.

16. CRESWELL, EDUCATIONAL RESEARCH, *supra* note 11, at 55–56.

participants' experiences—the experiences of both writers and readers.

As the name suggests, mixed methods research combines both quantitative and qualitative methods to gain the fullest possible understanding of the subject under investigation.<sup>17</sup> Mixed methods have become increasingly popular in other disciplines and offer great promise for the study of legal writing because they combine the strengths of quantitative and qualitative methodology.<sup>18</sup> Leading mixed methods researchers have even founded an international society focused on promoting mixed methods research across a variety of disciplines.<sup>19</sup>

Your choice of research method depends on the question you are investigating. If you are trying to discover which factors influence or predict an outcome, quantitative methods may be appropriate. If you are trying to understand a phenomenon, qualitative methods may serve you well. To answer both types of questions, you may pursue a mixed methods approach.<sup>20</sup>

The next three sections introduce common approaches to quantitative, qualitative, and mixed methods research. Although each section classifies the types of studies in each approach, please understand that no bright lines separate these categories. Think of them as points along a continuum, rather than exclusive classifications. In addition, terminology varies among different empirical researchers, so you may find texts or articles presenting different classification systems than you see below.

## II. QUANTITATIVE RESEARCH

### A. Basic Methodology of Quantitative Research

The following subsections explain four fundamental concepts in quantitative research. The first subsection explains populations and random samples. The second explains the role of different types of variables in quantitative research. The third explains

---

17. CRESWELL, RESEARCH DESIGN, *supra* note 5, at 4; PATTON, *supra* note 12, at 64–65; HANDBOOK OF MIXED METHODS IN SOCIAL AND BEHAVIORAL RESEARCH 8–12 (Abbas Tashakkori & Charles Teddlie eds., 2d ed. 2010).

18. CRESWELL, RESEARCH DESIGN, *supra* note 5, at 216–17; PATTON, *supra* note 12, at 89–92.

19. See MMIRA Mixed Methods International Research Ass'n, <http://mmira.wildapricot.org> (last visited June 18, 2015).

20. CRESWELL, RESEARCH DESIGN, *supra* note 5, at 20.

hypothesis testing and statistical significance. And the fourth distinguishes correlation from causation.

### 1. Populations and Samples

In quantitative analysis, the “population” is the entire group under study, whereas the “sample” is the subset drawn from the population and examined in some way.<sup>21</sup> Sampling makes it possible to study a manageable group (e.g., 100 lawyers) instead of the entire population (e.g., 1 million lawyers).<sup>22</sup>

Through statistical tools developed over several centuries of mathematics, researchers can draw inferences about the population based on a validly drawn sample.<sup>23</sup> This statistical “magic,” however, presumes that samples are randomly drawn from the population because random sampling gives researchers the best chance at drawing a sample representative of the population.<sup>24</sup>

Researchers can achieve random sampling in a variety of ways. *Simple random sampling* requires that every member of the population has an equal chance of inclusion.<sup>25</sup> For example, researchers could assign each member of the population a number and then use a random number table to choose which population members to survey.<sup>26</sup> *Stratified sampling* ensures that specific traits (e.g., gender) appear in the same proportions in the sample as they do in the population.<sup>27</sup> A researcher might choose stratified sampling where the “stratum of interest is a small percentage of a population,” in which case a simple random sample could easily underrepresent that stratum.<sup>28</sup> For example, a researcher studying a population of 20,000 college students may know that 2% (400) of the students are divorced parents with young children. If including this stratum in the sample is important, then the researcher may obtain a list of all 400 students in the

---

21. SMITH, *supra* note 11, at 142.

22. *Id.* at 142.

23. *Id.* at 164.

24. W. LAWRENCE NEUMAN, SOCIAL RESEARCH METHODS: QUALITATIVE AND QUANTITATIVE APPROACHES 218, 233 (5th ed. 2003).

25. *Id.* at 218; CRESWELL, RESEARCH DESIGN, *supra* note 5, at 158.

26. NEUMAN, *supra* note 24, at 218.

27. *Id.* at 223.

28. *Id.*

stratum, and select 4 students at random to include in the sample.<sup>29</sup>

## 2. Variables

Quantitative researchers use variables to quantify features of the phenomenon they are investigating.<sup>30</sup> *Dependent variables* are those that measure the outcome that researchers are investigating.<sup>31</sup> They are also known as output, response, or effect variables.<sup>32</sup> For example, in a study examining the correlation between note-taking methods and exam scores, the dependent variable would be the exam score.

*Independent variables* are those that researchers hypothesize may influence the outcome of interest.<sup>33</sup> Independent variables are also known as input variables, predictor variables, or treatment variables.<sup>34</sup> For example, in note-taking study, the independent variable might be each subject's note-taking method—laptop or longhand.<sup>35</sup>

*Control variables* are a special subset of independent variables. Although control variables may influence the outcome, researchers are not investigating these variables. Instead, researchers use control variables to help isolate the correlation between the independent variable and the outcome in question.<sup>36</sup> In the note-taking study, control variables might include demographic information (such as gender, age, and ethnicity) and incoming academic indicators (such as LSAT, UGPA, or law school GPA).

---

29. *Id.*

30. CRESWELL, RESEARCH DESIGN, *supra* note 5, at 25–53; NEUMAN, *supra* note 24, at 149. Researchers distinguish between variables and attributes. Variables are the broad categorizations, whereas attributes are the values or categories that the variable may take. For example, if “gender” was a variable in a given study, “male” and “female” would be attributes of that variable. NEUMAN, *supra* note 24, at 149. Similarly, if “years since bar admission” was a variable, the different possible years since bar admission for any given subject would be attributes of that variable. *See id.*

31. *Id.* at 149.

32. CRESWELL, RESEARCH DESIGN, *supra* note 5, at 52; NEUMAN, *supra* note 24, at 149.

33. NEUMAN, *supra* note 24, at 149.

34. CRESWELL, RESEARCH DESIGN, *supra* note 5, at 52.

35. This hypothetical is modeled roughly on Pam A. Mueller & Daniel M. Oppenheimer, *The Pen Is Mightier Than the Keyboard: Advantages of Longhand over Laptop Note Taking*, 25 PSYCHOL. SCI. 1159 (2014).

36. CRESWELL, RESEARCH DESIGN, *supra* note 5, at 53.

*Intervening variables* are those that may stand between the independent and dependent variables and mediate the effect of the independent variables on the dependent variable.<sup>37</sup> In our note-taking study, an intervening variable might be the student's exam preparation mode (e.g., whether or not the student organized his or her notes before studying).

Finally, *confounding variables* are those that the study does not measure, but that may nevertheless explain the relationship between the independent and dependent variables.<sup>38</sup> In the note-taking study, confounding variables might include students' study habits (other than note-taking style), degree of preparation (did the student prepare in advance or cram at the last minute), extra-curricular workload, or a personal crisis just before exam time.

### 3. *Hypothesis Testing and Statistical Significance*

In quantitative studies, the hypothesis is the researcher's prediction about how the variables in the sample will relate.<sup>39</sup> Quantitative research involves two types of hypotheses: the "null hypothesis" ( $H_0$ ) and the "alternative hypothesis" ( $H_1$ ).

The null hypothesis predicts that, within the population under study, no relationship exists between the independent and dependent variables.<sup>40</sup> The null hypothesis is often a "straw assumption" that the researchers expect to reject.<sup>41</sup> In our note-taking study, for example, the null hypothesis is that no relationship exists between the method of note-taking and exam scores.

Unsurprisingly, the alternative hypothesis assumes a relationship between the independent and dependent variables.<sup>42</sup> The alternative hypothesis may be directional or non-directional.<sup>43</sup> A directional hypothesis predicts the direction of the relationship between the independent and dependent variables, while a non-directional hypothesis merely predicts some type of relationship.<sup>44</sup> In our note-taking study, for example, a directional alternative

---

37. *Id.*; NEUMAN, *supra* note 24, at 150.

38. CRESWELL, RESEARCH DESIGN, *supra* note 5, at 53.

39. *Id.* at 143.

40. *Id.* at 144.

41. SMITH, *supra* note 11, at 192.

42. CRESWELL, RESEARCH DESIGN, *supra* note 5, at 144–45.

43. *Id.* Smith refers to these as "one-sided" and "two-sided" hypotheses. SMITH, *supra* note 11, at 192.

44. CRESWELL, RESEARCH DESIGN, *supra* note 5, at 144–45.



hypothesis might predict that longhand note-takers will receive higher test scores. In contrast, if prior studies offered no basis to predict whether longhand note-takers would receive higher scores, a non-directional alternative hypothesis would simply predict a significant difference between the exam scores of longhand and laptop note-takers.

Quantitative researchers test the null hypothesis by determining whether the result obtained is “statistically significant.”<sup>45</sup> Tests of statistical significance assume that no relationship exists between the variables of interest in the population (i.e., that the null hypothesis is true) and determine the likelihood that the relationship between the variables in the sample was the product of random chance in the sampling process.<sup>46</sup> Tests of statistical significance yield a “*p*-value,” a number between zero and one.<sup>47</sup> Although no fixed formula distinguishes significant from insignificant results, the most widely accepted significance level is .05.<sup>48</sup> In essence, this means that if there were no correlation between the two variables in the population, there would be only a five percent chance of seeing this correlation in the sample. Where the *p*-value is lower than .05, researchers report that as well—most commonly at the levels of  $p < 0.01$  or  $p < 0.001$ .<sup>49</sup> The smaller the *p*-value, the more confidence the researchers can have in rejecting the null hypothesis; however, rejecting the null hypothesis does not prove the alternative hypothesis to be true; it merely means that the researcher’s results are *consistent with* the alternative hypothesis.<sup>50</sup>

By way of illustration, if the note-taking study found that the longhand note-takers had a higher average exam score than laptop note-takers and if the difference yielded a significance level of  $p < .05$ , then researchers could reject the null hypothesis that note-taking methodology bears no relationship to exam score. On

---

45. NEUMAN, *supra* note 24, at 356.

46. *Id.* at 356–57.

47. NEUMAN, *supra* note 24, at 356–57; SMITH, *supra* note 11, at 192–96.

48. NEUMAN, *supra* note 24, at 357; SMITH, *supra* note 11, at 196. Researchers have simply adopted R.A. Fisher’s arbitrary adoption of 0.05 as the cutoff for statistical significance, although some researchers treat 0.05 as “significant” and 0.01 as “highly significant.” SMITH, *supra* note 11, at 196.

49. NEUMAN, *supra* note 24, at 357.

50. *Id.* at 150–51.

the other hand, if the study yielded a significance level of .15, researchers could not reject the null hypothesis.

#### 4. *Correlation Versus Causation*

In any analysis that involves correlation, the researcher must remember a fundamental tenet of statistical analysis: correlation is not causation.<sup>51</sup> One colorful illustration of this tenet is the so-called Super Bowl Indicator “discovered” in 1978 by *New York Times* sportswriter Leonard Koppett.<sup>52</sup> This tongue-in-cheek theory holds that the stock market goes up if an NFC team or a former NFL team currently in the AFC wins the Super Bowl.<sup>53</sup> Prior to the 2014 Super Bowl, the indicator had been correct in 35 out of 47 years (from 1967 to 2013).<sup>54</sup> Assuming there was no relationship between Super Bowl victors and the stock market, the likelihood of that correlation occurring through random chance was only 2.9 percent, which would yield a statistically significant *p*-value of 0.029.<sup>55</sup> Common sense, of course, dictates that no plausible theory could explain why the Super Bowl victor would affect the stock market. The unlikely correlation is simply a result of random chance.

An unexpected correlation could still signal a relationship worth exploring. But the researcher must make sure that a plausible, common-sense theory can explain how the variables might interact. Mixed methods research could be a way to develop or explore such a theory. Alternatively, after finding a correlation of

---

51. SMITH, *supra* note 11, at 261–62. Smith discusses three reasons why a correlation might exist despite the lack of causation: “simple chance, reverse causation, and omitted factors.” *Id.* at 262.

52. *Id.* at 205; Robert R. Johnson, *Is It Time to Sack the ‘Super Bowl Indicator?’*, WALL STREET J., Total Return Blog, Jan. 22, 2014, available at <http://blogs.wsj.com/totalreturn/2014/01/22/is-it-time-to-sack-the-super-bowl-indicator/>. Smith also refers to a correlations between “the stock market and the number of strikeouts by a professional baseball team,” and between “the annual rate of inflation and the number of dysentery cases in Scotland the previous year.” SMITH, *supra* note 11, at 262; see also *Spurious Correlations*, SPURIOUS MEDIA LLC, [www.tylervigen.com](http://www.tylervigen.com) (last visited June 15, 2015). The website of Tyler Vigen, “Spurious Correlations,” identified many improbable correlations, including a 0.992558 correlation between the divorce rate in Maine and the per capita consumption of margarine 2000 to 2009. *Id.*

53. SMITH, *supra* note 11, at 205.

54. *Id.*

55. Johnson, *supra* note 50.

interest, a researcher could develop an experimental design to test whether a causal relationship exists.

## B. Quantitative Strategies of Investigation

Quantitative methods can be experimental or non-experimental.<sup>56</sup> Experimental designs manipulate the world you are studying, while non-experimental designs merely study the world as you find it.<sup>57</sup> This section will summarize key features of the most common non-experimental and experimental designs and illustrate each design by describing empirical studies of legal writing.

### 1. *Non-Experimental Designs*

Quantitative researchers can gather data directly or indirectly. Researchers gather data directly from their subjects through surveys and assessments. Researchers gather data indirectly by analyzing the content of the subjects' writing.

#### a. Survey and Assessment Designs

The following subsections describe the features of survey and assessment design and use several empirical legal writing studies to illustrate those features.

##### *i. Features of Surveys and Assessments*

In a survey, researchers try to learn about trends, attitudes, opinions, and practices in a population by surveying a sample of that population.<sup>58</sup> Researchers use a variety of survey methodologies. They may survey participants directly, by stopping people in public places, going door-to-door, or calling them on the telephone. Alternatively, researchers may mail or email instruments for the participants to complete themselves.<sup>59</sup>

---

56. PAUL C. COZBY & SCOTT C. BATES, *METHODS IN BEHAVIORAL RESEARCH* ch. 4 (11th ed. 2011).

57. *Id.*

58. CRESWELL, *RESEARCH DESIGN*, *supra* note 5, at 157; NEUMAN, *supra* note 24, at 233.

59. NEUMAN, *supra* note 24, at 289–92, 300–01 (discussing advantages, disadvantages, and costs of various survey methodologies); *see* CRESWELL, *RESEARCH DESIGN*, *supra* note 5, at 157.

Researchers quantify their results in a variety of ways, depending on what the survey instrument asked. They may ask participants to select among categories (e.g., which religion you practice, if any; which political candidate respondents prefer); or they may call for responses on numerical scales (e.g., rating agreement with a series of statements on a scale of 1-5).<sup>60</sup>

In addition to asking people about their opinions or beliefs, researchers can use instruments to assess people's knowledge, skills, or practices. In some cases, researchers can access assessments that the subjects have already taken, such as the LSAT.<sup>61</sup> In other cases, researchers may administer assessment instruments as part of their study, as in Marjorie Shultz and Sheldon Zedek's work on predicting successful lawyering.<sup>62</sup>

Samples must be random in order for researchers to infer population characteristics from the sample characteristics.<sup>63</sup> In addition, researchers must ensure that they have a sufficient sample size to draw inferences about the population.<sup>64</sup> For surveys, the necessary sample size depends not only on the population size, but also on the degree of accuracy required (e.g., 95% confidence, 99% confidence), the degree of variability expected in the population, and the number of variables to be analyzed simultaneously.<sup>65</sup> For researchers unable to estimate the parameters needed to determine the sample size, a useful rule of thumb is that smaller populations require sampling a larger percentage of the population.<sup>66</sup> This is because as the size of a population grows, the returns in accuracy from increased sample

---

60. NEUMAN, *supra* note 24, at 188, 195–96.

61. See, e.g., David A. Thomas, *Predicting Law School Academic Performance from LSAT Scores and Undergraduate Grade Point Averages: A Comprehensive Study*, 35 ARIZ. ST. L.J. 1007 (2003).

62. Marjorie M. Shultz & Sheldon Zedek, *Identification, Development, and Validation of Predictors for Successful Lawyering* (Sept. 2008), available at <http://ssrn.com/abstract=1353554>.

63. See *supra* text at nn. 21–29.

64. NEUMAN, *supra* note 24, at 232.

65. *Id.* To get a feel for how the degrees of certainty (“confidence level”) and margin of error (“confidence interval”) affect the necessary sample size, you may experiment with a free online calculator available at [www.surveysystem.com/sscalc.htm](http://www.surveysystem.com/sscalc.htm). Beyond a certain point, increasing the population size has almost no effect on the necessary sample size. But very small changes to the confidence level and confidence interval can have dramatic effects.

66. NEUMAN, *supra* note 24, at 232.

size shrink.<sup>67</sup> So a population of 1,000 might require a sample of 300 (30%), whereas a population of 150,000 might require a sample of 1,000 to 1,500 (1% to 1.5%).<sup>68</sup>

ii. *Illustrations of Surveys and Assessments*

Robert Benson and Joan Kessler's study illustrates how researchers can compare the responses of two different groups.<sup>69</sup> Benson and Kessler studied whether "legalese" writing affected judges' or clerks' opinions of (1) a brief's persuasiveness, and (2) the author's credibility.<sup>70</sup> The sample was thirty-three law clerks and ten appellate judges from one California appellate district.<sup>71</sup>

The researchers presented the participants with short excerpts from an appellate brief and a petition for rehearing.<sup>72</sup> Some participants received "legalese" versions laden with complex sentences, jargon, long words, extra words, and nominalizations.<sup>73</sup> Others received streamlined "plain language" versions.<sup>74</sup> The survey asked participants to rate their disagreement or agreement (on a 1–5 scale) with a number of statements about the brief's content and persuasiveness, as well as the author's qualifications, professional credentials, and personal credibility.<sup>75</sup>

The legalese version received significantly lower assessments of the brief's persuasiveness and of the author's qualifications and professional credentials (though not of the author's personal credibility).<sup>76</sup> Readers of the legalese briefs were significantly more likely to agree that the author was "unscholarly," worked at a less prestigious firm, and was not an effective appellate advocate.<sup>77</sup>

---

67. *Id.*

68. *Id.*

69. Robert W. Benson & Joan B. Kessler, *Legalese v. Plain English: An Empirical Study of Persuasion and Credibility in Appellate Brief Writing*, 20 LOY. L.A. L. REV. 301 (1987).

70. *Id.* at 301–02.

71. *Id.* at 305.

72. *Id.* at 306.

73. *Id.* at 307–11.

74. *Id.*

75. *Id.* at 312–13.

76. *Id.* at 313–15.

77. *Id.* For those statements, the difference between responses of those who read the legalese and those who read the plain language were statistically significant at either  $p = .01$  or  $p = .001$ . *Id.*

Sean Flammer's study took a somewhat different approach to studying how judges respond to legalese.<sup>78</sup> Flammer drafted three versions of the same two-page excerpt from a brief. One was a legalese version, which suffered from the same flaws as Benson and Kessler's legalese excerpt.<sup>79</sup> The second was a plain language version.<sup>80</sup> The third was an informal version that adopted a conversational style and included contractions and the first person.<sup>81</sup>

Flammer sent surveys to federal and state trial and appellate judges.<sup>82</sup> Half received the legalese and the plain English versions; the other half received the legalese and informal versions.<sup>83</sup> Flammer asked each participant which version "was more likely to persuade you."<sup>84</sup>

Overall, the judges preferred plain English to legalese by 66% to 34%.<sup>85</sup> This preference was largely consistent across different categories based on gender, age, and years on the bench.<sup>86</sup> The judges also preferred the informal to the legalese by 58% to 42%, a somewhat smaller preference.<sup>87</sup> The only variable showing a strong correlation to the legalese versus informal preference was gender. While 83% of female judges preferred the informal brief to the legalese brief, only 51% of male judges preferred the informal brief to the legalese brief.<sup>88</sup>

These studies illustrate how researchers used surveys to investigate whether writing style affects whether readers view the writing as persuasive. Both studies designed survey instruments that they mailed to the participants. And both studies explored readers' perceptions by randomly assigning participants to read one of several different writing samples and rate the sample's persuasiveness.

---

78. Sean Flammer, *Persuading Judges: An Empirical Analysis of Writing Style, Persuasion, and the Use of Plain English*, 16 *LEGAL WRITING* 183 (2010).

79. *Id.* at 193–95.

80. *Id.*

81. *Id.* at 195–97.

82. *Id.* at 190–91.

83. *Id.*

84. *Id.* at 220.

85. *Id.* at 201.

86. *Id.*

87. *Id.* at 206.

88. *Id.* at 208. Joseph Kimble also conducted a series of surveys examining lawyers' and judges' preferences for plain English over legalese in briefs, contracts, and statutes. See JOSEPH KIMBLE, *LIFTING THE FOG OF LEGALESE: ESSAYS ON PLAIN LANGUAGE* 3–13 (2006).

b. Content Analysis

i. *Features of Content Analysis*

In addition to surveying people, quantitative researchers can also gather information by coding existing documents. Lawrence Neuman refers to this as content analysis, a subset of non-experimental quantitative research design.<sup>89</sup> In content analysis, researchers code and analyze aspects of written or spoken texts, and compare content across many texts.<sup>90</sup> Like surveys, content analysis involves random sampling.<sup>91</sup>

Content analysis is well suited to “problems involving a large volume of text,”<sup>92</sup> as may be the case in a study of legal writing. Researchers can study large numbers of documents by implementing well-constructed codes, careful protocols, and multiple coders or automated coding tools.<sup>93</sup> Researchers determine the relevant “unit of analysis”<sup>94</sup> depending on the research question. For example, researchers may study narrow units like specific words or phrases, or broader units like theme or plot.<sup>95</sup> “Manifest coding” examines content that is evident from the text itself, without any interpretation by the researcher.<sup>96</sup> In legal writing studies, manifest coding might examine the use of particular words, word counts, or readability scores. In contrast, “latent coding” examines implicit meanings or themes.<sup>97</sup> In legal writing studies, latent coding might examine themes or narrative elements.

Some content analysis studies are purely descriptive rather than correlational.<sup>98</sup> Those studies present information about the

---

89. NEUMAN, *supra* note 24, at 310. Neuman distinguishes between reactive and non-reactive designs, instead of experimental and nonexperimental. *Id.* at 308. Surveys and experiments are reactive because participants are aware that they are being studied. *Id.* Content analysis is nonreactive because the participants are unaware that they are research subjects. *Id.*

90. *Id.* at 311.

91. *Id.*

92. *Id.*

93. *Id.* at 311–12.

94. *Id.* at 312.

95. *Id.*

96. *Id.* at 313.

97. *Id.*

98. DANIEL RIFFE ET AL., *ANALYZING MEDIA MESSAGES: USING QUANTITATIVE CONTENT ANALYSIS IN RESEARCH* 19, 26–28 (3d ed. 2014).

texts under study but do not analyze the relationship among the variables.<sup>99</sup> In contrast, other content analysis studies are correlational because they analyze the extent to which variables are related to one another.<sup>100</sup> Correlation exists when a change in one variable is associated with a change in another.<sup>101</sup> When the change moves both variables in the same direction, the correlation is positive.<sup>102</sup> When the variables move in the opposite direction, the correlation is negative.<sup>103</sup> For example, if every one-point increase in a student's LSAT score was associated with a 0.1-point increase in the student's law school GPA, then researchers would describe LSAT and law school GPA as positively correlated.

*ii. Illustrations of Content Analysis—Descriptive Studies*

In a descriptive content analysis study, researchers analyze the characteristics of, for example, a set of documents. In one comprehensive example, Brady Coleman and his colleagues studied over 2,500 questions presented in Supreme Court briefs filed from 1953 to 2002.<sup>104</sup> To collect data from the briefs, the researchers developed a computer program to extract the question presented from each brief and measure several variables: number of questions presented, words per question presented, use of an introductory statement, numbering style, first word, and ending punctuation.<sup>105</sup>

In addition to reporting their overall findings for each variable, the researchers reported how the Office of the Solicitor

---

99. *See id.*

100. KLAUS KRIPPENDORFF, *CONTENT ANALYSIS: AN INTRODUCTION TO ITS METHODOLOGY* 192–94 (3d ed. 2013); *see also* RIFFE ET AL., *supra* note 98, at 27–28 (discussing use of content analysis along with use of statistical tools to infer meaning from the observed data).

101. NEUMAN, *supra* note 24, at 348–49.

102. *Id.*

103. *Id.* Note that the strength of the correlation is unrelated to the size of the change. For example, researchers might find that every additional week of study reliably associates with one-tenth of one percent increase in exam score. In that case, the correlation coefficient would be near-perfect, but the size of the change in exam score would be extremely small. *See* WILLIAM MENDENHALL ET AL., *INTRODUCTION TO PROBABILITY AND STATISTICS* 513–16 (14th ed. 2013) (explaining the formula for the Pearson Product Moment Coefficient of Correlation, which proves this principle).

104. Brady S. Coleman et al., *Grammatical and Structural Choices in Issue Framing: A Quantitative Analysis of “Questions Presented” from a Half Century of Supreme Court Briefs*, 29 AM. J. TRIAL ADVOC. 327, 330 (2005).

105. *Id.* at 331.



General (“OSG”) briefs differed from other briefs.<sup>106</sup> For example, the number of questions per brief was 1.78 for OSG briefs, and 2.58 for non-OSG briefs.<sup>107</sup> Similarly, the average words per question was 46.58 in petitioner briefs and 44.31 in respondent briefs, and this petitioner-respondent gap was larger in OSG briefs.<sup>108</sup> The researchers also reported on trends over time. For example, the average number of questions per brief declined from 1975 to 2002.<sup>109</sup>

Michael Murray studied rhetorical uses of parentheticals in federal appellate briefs and opinions.<sup>110</sup> He drew a cross-sectional sample of briefs filed in several appellate courts from February through July of 2011. The sample include fifty briefs filed from February to July of 2011 in the Supreme Court and three circuit courts, as well as opinions from fifty Supreme Court cases in the same time period.<sup>111</sup>

Murray coded the briefs for several different ways lawyers use cases in legal argument. Quotation parentheticals quote or highlight the content of the cited authority.<sup>112</sup> In contrast, “explanatory parentheticals” communicate the “lessons and principles induced from a synthesis of authorities.”<sup>113</sup> Murray also identified two subsets of explanatory parentheticals. “Public policy synthesis” parentheticals “explain or demonstrate the operation of public policy within multiple authorities.”<sup>114</sup> And “narrative synthesis” parentheticals “communicate narrative reasoning from the storylines of multiple authorities.”<sup>115</sup> Murray’s coding also identified two other ways of explaining authorities: case-by-case analogies in the text and descriptions in footnotes.<sup>116</sup>

---

106. *Id.* at 332.

107. *Id.* at 346.

108. *Id.* at 340–41.

109. *Id.* at 347.

110. Michael Murray, *The Promise of Parentheticals: An Empirical Study of the Use of Parentheticals in Federal Appellate Briefs*, 10 *LEGAL COMM. & RHETORIC: JALWD* 229 (2013) (using the term “rhetorical” to exclude parentheticals conveying purely bibliographical or procedural information).

111. *Id.* at 252–54.

112. *Id.* at 232–33.

113. *Id.* at 231.

114. *Id.*

115. *Id.*

116. *Id.* at 250–52.

Murray found that practitioners use explanatory parentheticals more frequently than the other rhetorical devices.<sup>117</sup> He also found that parentheticals were the most common way for practitioners to present analogical comparison of multiple authorities.<sup>118</sup> Next, Murray found that each instance of explanatory synthesis parentheticals presented, on average, more authorities than each instance of case-to-case analogical reasoning.<sup>119</sup> Finally, Murray found that these two methods are not mutually exclusive; the vast majority of briefs used both explanatory synthesis parentheticals and case-to-case analogical reasoning.<sup>120</sup>

These studies illustrate how researchers can analyze the content of legal writing with descriptive studies. In both studies, the researchers coded several variables that described the content of the writing. Coleman and his colleagues used manifest coding to identify word counts, numbering style, and punctuation. Murray, in contrast, used latent coding to classify the authors' use of parentheticals. Both studies were descriptive because their goal was to identify certain features of the writing rather than examine the relationship among variables.

### *iii. Illustrations of Content Analysis: Correlational Studies*

Unlike descriptive studies, correlational studies analyze the relationship between variables. Paul Collins, Jr., Pamela Corley, and Jesse Hamner studied the influence of U.S. Supreme Court amicus briefs by analyzing over 2,000 amicus briefs and over 200 opinions from the U.S. Supreme Court's 2002–2004 terms.<sup>121</sup> The dependent variable used to measure influence was the percentage of the majority opinion based directly on the amicus brief's language. The researchers used plagiarism detection software (WCopyfind) to identify what percentage of the Court's opinion was based on the brief.<sup>122</sup>

---

117. *Id.* at 255.

118. *Id.*

119. *Id.* at 256.

120. *Id.* at 258.

121. Pamela Corley et al., *The Influence of Amicus Curiae Briefs on U.S. Supreme Court Opinion Content* 11 (unpublished manuscript delivered at the 109th Annual Meeting of the American Political Science Ass'n, Chicago, Ill. Aug. 31 2013), available at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2300505](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2300505).

122. *Id.*

The independent variables measured in the amicus brief included clarity and plain language (scored with the content analysis tool Linguistic Inquiry and Word Count); repetition of the lower court opinion or other briefs (measured with WCopyfind); ideological congruence with the majority opinion writer; and whether the Solicitor General authored the brief.<sup>123</sup> The researchers also measured several control variables. They measured “case salience” based on news reports about the case. They recorded brief length and the number of amicus briefs filed in the case. And they coded the percentage of the majority opinion drawn from other briefs and opinions (measured by the plagiarism detection software).<sup>124</sup>

Collins and his colleagues found statistically significant ( $p < .01$ ) correlations between every independent and control variable and the dependent variable (percentage of the majority opinion based on the amicus brief). In addition, the variables were signed in the direction that the researchers predicted. For example, the researchers predicted that the correlation between ideological congruence and amicus brief adoption would be positively signed, meaning that a higher ideological congruence score would correlate with greater adoption of the amicus brief. Conversely, the researchers predicted that the correlation between the number of amicus briefs filed and amicus brief adoption would be negatively signed, meaning that a higher total number of amicus briefs in a given case would correlate with less adoption of the amicus brief in question. This study is especially noteworthy for its use of automated content analysis software to study the content of over 2,000 briefs and opinions.<sup>125</sup>

In another correlational study, Lance Long and William Christenson studied the relationship between readability and outcome in appellate briefs.<sup>126</sup> They sampled 882 appellate briefs from the Supreme Court, federal appellate courts, and state supreme courts. Their dependent variable was the outcome of the appeal (affirmed or reversed). Their independent variable was readability, as measured by the Flesch Reading Ease score

---

123. *Id.* at 14–17.

124. *Id.*

125. *Id.* at 18–21.

126. Lance Long & William F. Christensen, *Does the Readability of Your Brief Affect Your Chance of Winning an Appeal?*, 12 J. APP. PRAC. & PROCESS 145 (2011).

calculated by Microsoft Word.<sup>127</sup> For federal appellate and state supreme court briefs, the researchers coded control variables for federal or state court, standard of review, presence of a dissenting opinion (present or absent), and readability of the opinion deciding the appeal. For U.S. Supreme Court briefs, the researchers coded control variables for constitutional issue, criminal or civil case, presence of a dissenting opinion, and opinion readability.<sup>128</sup>

Long and Christensen found no statistically significant correlation between readability and outcome in the briefs in their study.<sup>129</sup> For federal appellate court briefs and state supreme court briefs, the only variable with a statistically significant correlation to reversal was jurisdiction, which is not surprising because state cases have higher reversal rates than federal cases. For U.S. Supreme Court briefs, no variable showed a statistically significant correlation with reversal.<sup>130</sup>

The researchers did, however, find an interesting relationship between the readability of briefs and opinions. First, readability scores for federal appellate court opinions and state supreme court opinions were similar.<sup>131</sup> Second, readability scores from U.S. Supreme Court opinions were lower than opinions from the other courts.<sup>132</sup> Third, briefs from U.S. Supreme Court cases had lower readability scores than briefs from the other courts.<sup>133</sup> Finally, at the U.S. Supreme Court level, the opinion readability scores were significantly lower than the brief readability scores.<sup>134</sup>

These studies illustrate how researchers use correlational studies to explore the relationship between variables. Long and Christensen analyzed the relationship between readability and outcome on appeal, whereas Collins and his colleagues analyzed the relationship between various writing-based and author-based

---

127. *Id.* at 150–51. The FRE score is a function of the average number of words syllables per word and words per sentence. *Id.* at 151.

128. *Id.* at 155.

129. *Id.* at 156–57. Although this study found no significant correlation between readability and outcome, that could reflect lower caseloads and greater law clerk support for appellate judges in comparison to trial judges. For that reason, I have launched a readability study of a different population—state and federal summary judgment briefs. My study will also include a different readability measure—a differential score—which indicates the difference between each brief's readability score and the score of its opposite number.

130. *Id.* at 157–58.

131. *Id.* at 157.

132. *Id.*

133. *Id.* at 157–58.

134. *Id.*

variables and the degree to which the U.S. Supreme Court opinion adopted the language of amicus briefs.

## 2. *Experimental Designs*

The following subsections identify the core features of experimental designs and illustrate these features through several legal writing studies.

### a. Features of Experimental Designs

In a *true experiment*, researchers randomly assign participants to groups and manipulate the conditions so that one group receives the treatment (the treatment group) and the other does not (the control group).<sup>135</sup> Random assignment to treatment and control groups minimizes the likelihood of confounding variables.<sup>136</sup> Confounding variables are more prevalent in non-experimental studies in which researchers cannot manipulate exposure to the treatment. That is why experiments offer stronger evidence of causation than non-experimental designs.<sup>137</sup>

When a study is similar to an experiment but lacks random assignment, researchers commonly refer to it as a *quasi-experiment*.<sup>138</sup> Random assignment is often impossible when researchers must study naturally-formed groups such as classrooms or families. Researchers refer to the samples in such cases as convenience samples.<sup>139</sup>

Researchers analyze the data by using statistical tools to test their hypotheses. These statistical tools evaluate the relationships among variables and groups and test for statistically significant differences in the outcomes of the treatment and control groups. Once their statistical analyses are complete, researchers interpret the data by theorizing about why the treatment did or did not make a difference.

Researchers must be wary of threats to the validity of their experiments. These threats may be internal or external to the

---

135. CRESWELL, RESEARCH DESIGN, *supra* note 5, at 168, 170.

136. NEUMAN, *supra* note 24, at 240.

137. *Id.* at 238.

138. CRESWELL, RESEARCH DESIGN, *supra* note 5, at 168.

139. *Id.*

experiment. *Internal threats* threaten “the researcher’s ability to draw correct inferences from the [sample] data about the population.”<sup>140</sup> For example, participants may change or mature during the course of the experiment, participants may drop out during the experiment, or participants in the treatment and control groups could communicate with each other and influence the outcomes.<sup>141</sup> *External threats* threaten the researcher’s ability to draw inferences from the sample data to other populations or contexts.<sup>142</sup> For example, the results may not generalize beyond the social groups represented in the study, or may not generalize to settings not under study or to past or future situations.<sup>143</sup>

### b. Illustrations of Experimental Designs

So far no legal writing researchers appear to have published what one might call a “classic” experiment, which involves not only random assignment of participants to treatment and control groups, but also a pre-test and a post-test.<sup>144</sup> However, Kenneth Chestek conducted a “post-test-only” experiment, which means that the only measurement of the dependent variable took place after the treatment and control groups were exposed to different versions of the independent variable.<sup>145</sup> Chestek studied whether judges’ assessments of briefs are influenced by narratives that litigants present.<sup>146</sup> Chestek created a fictional case that was a “hard case” for the plaintiff but an “easy one” for the defendant.<sup>147</sup> Chestek also created two briefs for each side: a “logos” brief and a “story” brief. The logos brief was a straightforward presentation of the critical facts and legal arguments. In contrast, the story brief wove a compelling narrative for each side, in addition to

---

140. *Id.* at 174.

141. *Id.* at 174–75.

142. *Id.* at 176. For a discussion of internal and external validity threats and how to avoid them, see *id.* at 174–77.

143. *Id.* at 176.

144. *Id.* at 168, 173.

145. *Id.* at 173; NEUMAN, *supra* note 24, at 247. In a post-test-only design, the random assignment “reduces the chance that the groups differed before the treatment, but without a pretest, a researcher cannot be as certain that the groups began the same on the dependent variable.” NEUMAN, *supra* note 24, at 247.

146. Kenneth D. Chestek, *Judging by the Numbers: An Empirical Study of the Power of Story*, 7 J. ALWD 1 (2010). The ninety-five survey respondents were appellate judges, staff attorneys, clerks, and practitioners, as well as law professors. *Id.*

147. *Id.* at 10–11.

presenting the critical facts and legal arguments. To isolate the role of the narrative, both the logos brief and the story brief relied on essentially the same statement of the issue, authorities, and legal argument, although the arguments obviously varied based on which party was arguing.<sup>148</sup>

Chestek gave each participant the logos brief and story brief for the same side of the fictional case. Some read the “easy” side, and the rest read the “hard” side. Participants read briefs for the same side in order to isolate the role of story as a variable and to avoid having the merits of the underlying case skew the participants’ assessments of each brief’s persuasiveness. Chestek then asked the participants to rate the persuasiveness of each brief on a scale from 1 to 5, with 5 being most persuasive.<sup>149</sup>

The participants rated the story brief more persuasive than the logos brief, which confirmed Chestek’s hypothesis that briefs with a strong narrative component are more persuasive than purely law-based briefs. However, the degree of preference was not significantly different for the “easy” side compared to the “hard” side.<sup>150</sup> This finding refuted Chestek’s hypothesis that the persuasiveness gap between the story brief and the logos brief would be larger in hard cases than in easy cases.<sup>151</sup>

Chestek’s findings also suggest that a lawyer’s experience and role in the system may affect his or her perspective on what is persuasive. More experienced judges and attorneys were more likely to find the story brief more persuasive, whereas law clerks were evenly divided on which was more persuasive.<sup>152</sup> In addition, practitioners expressed the strongest preference for the story briefs, which is consistent with their primary focus on representing clients.<sup>153</sup> In contrast, judges and law clerks expressed the

---

148. *Id.* at 11–14.

149. *Id.* at 18–19. Because he used a survey instrument to measure the dependent variable (persuasiveness), Chestek conducted a “population-based survey experiment.” For a discussion of population-based survey experiments, see DIANA C. MUTZ, POPULATION-BASED SURVEY EXPERIMENTS 2 (2011). In a population-based survey experiment, the researcher (1) uses survey sampling methodology to produce a sample representative of the population of interest, and (2) randomly assigns “participants to variations of the independent variable in order to observe their effects on a dependent variable.” *Id.*

150. Chestek, *supra* note 146, at 18–19.

151. *Id.*

152. *Id.* at 20.

153. *Id.*

lowest preference for the story briefs, perhaps because of the primacy of law in their institutional roles.<sup>154</sup>

### III. QUALITATIVE RESEARCH

The sections that follow discuss some fundamentals of qualitative research and then describe the key features and illustrations of five common qualitative research strategies.

#### A. Basic Methodology of Qualitative Research

The subsections that follow explain four core concepts in qualitative research. The first subsection explains the role that theory plays in qualitative research. The second explains basic qualitative data collection techniques. The third explains how qualitative researchers move from raw data to theory. Finally, the fourth explains the importance of theory verification in qualitative research.

##### 1. *The Role of Theory*

The role of theory in qualitative research often differs from the role of a hypothesis in quantitative research. In many studies, the researcher's goal is to develop a new theory inductively, based on the qualitative data gathered from the subjects. These researchers begin by gathering data from their subjects and construct their theory from that data. In some cases, however, the qualitative researcher's goal is to test, modify, or explore an existing theory about process, phenomenon, or people. In these cases, the researcher's theory plays a role similar to a quantitative

---

154. *Id.* at 20. Chestek's study could be modified in an interesting way to study whether narrative affects how judges decide a hypothetical case, rather than how judges rate persuasiveness. Researchers would randomly assign participants to one of four different groups: (1) a control group that receives the logos brief for each side; (2) a treatment group that receives the plaintiff logos and defendant story brief; (3) a treatment group that receives the plaintiff story and defendant logos brief; and (4) a treatment group that receives the story brief for each side. The participants would decide the hypothetical case. The researcher would analyze the data to test the null hypothesis that story briefs do not affect on judges' decisions. If the data showed a statistically significant increase in decisions for the story brief, the findings would support the alternative hypothesis that story briefs affect judges' decisions.



ypothesis because the theory frames the qualitative study in the same way that the hypothesis frames a quantitative study.<sup>155</sup>

## 2. *Qualitative Data Collection Techniques*

Unlike the instrument-based collection of quantitative data, qualitative researchers collect raw data primarily through observations, interviews, and document review. Observations often involve field notes on the subject's behavior in context. The researcher may range from a completely detached observer to a full participant in the subject's experience. Researchers may conduct individual interviews face-to-face or by phone. They may also interview multiple participants simultaneously through focus groups.<sup>156</sup> Interviews usually involve open-ended questions based on protocols that include standard prompts to ensure that the interviews cover the same core set of topics.<sup>157</sup>

## 3. *Moving from Raw Data to Theory*

Once the researcher has gathered all of the raw data, much work remains. The researcher must construct a theory from the raw data. What follows is John Creswell's approach to that process, though there are as many possible ways to tackle that task as there are researchers.

First, the researcher gains a general sense of the data by reading everything once and jotting down ideas. The researcher then reads a few documents carefully in search of deeper meaning and records thoughts about that more careful read.<sup>158</sup>

Next, the researcher lists the topics emerging from his or her notes.<sup>159</sup> The researcher clusters similar topics together and identifies major topics, unique topics, and "leftover" topics.<sup>160</sup> At this stage, the researcher should be on the lookout for not only the topics he or she expected to find but also surprising topics and topics that are of conceptual interest to the reader.<sup>161</sup> The

---

155. CRESWELL, RESEARCH DESIGN, *supra* note 5, at 64–67.

156. *Id.* at 190–91.

157. *Id.* at 194.

158. *Id.* at 197.

159. *Id.* at 198.

160. *Id.*

161. *Id.* at 198–99.

researcher creates abbreviations for the topics and uses them as the codes in subsequent steps.<sup>162</sup>

Third, the researcher delves back into the data and codes every passage that relates to the codes he or she developed in the prior step. The same content may relate to more than one code.<sup>163</sup>

Fourth, the researcher analyzes all of the data within each code and performs a preliminary analysis of each code. This analysis may reveal ambiguities, overlaps, or gaps in the codes, and the researcher may recode the data if necessary.<sup>164</sup>

Finally, the researcher uses the coding analysis in several ways. The researcher develops a detailed description of the people, events, or phenomena the researcher is studying.<sup>165</sup> The researcher also generates perhaps five to seven themes or categories which constitute the “major findings” of the study.<sup>166</sup> Current or former legal practitioners may find the process of moving from raw data to codes and themes similar to the coding of large sets of documents during discovery.

#### 4. *The Importance of Verification*

One challenge of working with qualitative data is the subjectivity inherent in the process, both from the perspective of the subjects providing the data and the researchers gathering and interpreting it. Qualitative researchers build in processes to ensure that their observations are reliable and valid.<sup>167</sup>

Qualitative observations are reliable when the data gathering process is consistent across researchers.<sup>168</sup> For example, if researchers placed an apple on five different scales, and each scale registered 50 pounds, that consistent measurement would be reliable, although probably not valid. One threat to reliability in qualitative research is “code drift.”<sup>169</sup> Over the course of a coding process, individual researchers’ understandings of what particular codes mean may begin to evolve. To guard against code

---

162. *See id.*

163. *Id.* at 198.

164. *See id.*

165. *Id.* at 199.

166. *Id.*

167. *Id.* at 201.

168. *Id.*

169. *Id.* at 203.

drift, researchers can conduct frequent researcher meetings to discuss their understandings of the code.<sup>170</sup> Researchers should also share their coding questions with each other early and often.<sup>171</sup>

Qualitative observations are valid when they accurately reflect both the participants' views and the researchers' views of the experience.<sup>172</sup> For example, if five researchers coded the same passage as displaying hostility, but the author's experience in writing the passage was one of insecurity, the coding would be reliable but not valid. One way to promote validity is through triangulation, which involves collecting multiple sources of data to study the same phenomenon.<sup>173</sup> Researchers can also promote validity by conducting peer debriefings among themselves, as well as member debriefings to get the subjects' reactions to the developing interpretations and theories.<sup>174</sup> In the example of researchers miscoding the author's text as hostile, member debriefing would probably reveal the coding as invalid.

## B. Qualitative Strategies of Investigation

The following subsections describe and illustrate five different types of qualitative strategies: case studies, grounded theory, phenomenological studies, ethnographic research, and narrative research.

### 1. Case Studies

In a case study, the researcher explores an issue in depth by studying one or more instances within a specific setting.<sup>175</sup> The issue under study may be an event, process, activity, or set of individuals. A case study involves detailed data collection over a sustained period through a variety of methods.<sup>176</sup> The researcher may report on findings unique to a particular case or may report themes that emerge from multiple cases. The report will involve

---

170. *Id.*

171. *Id.* at 202–03.

172. *Id.* at 201.

173. *Id.*

174. *Id.* at 201–02.

175. JOHN W. CRESWELL, *QUALITATIVE INQUIRY & RESEARCH DESIGN: CHOOSING AMONG FIVE APPROACHES* 73 (2d ed. 2007) [hereinafter CRESWELL, *QUALITATIVE INQUIRY*].

176. *Id.* at 74.

detailed description of the case or cases, a discussion of a few key issues specific to those cases, and then analysis of themes that transcend the cases.<sup>177</sup>

Chestek conducted another case study, this one on the use of narratives in litigation challenging the Affordable Care Act.<sup>178</sup> He analyzed a dozen district court cases and identified three different types of plaintiffs: (1) private individual and employers health-care consumers; (2) physicians and physician groups; and (3) state governments. The defendants in every case were various federal government officials. Chestek studied the different narratives presented by each category of litigant and explored how the different narratives interacted.<sup>179</sup>

For example, he noted that nearly all of the private plaintiffs lost on the issue of whether the ACA exceeded Congress's Commerce Clause power, but the state government plaintiffs won on that issue. He noted that the private plaintiffs told stories of "rugged individualism" that failed in the face of the federal government's story of Congress representing "everyperson." Because even rugged individuals need health insurance eventually, the rugged individuals contributed to the freeloading problem that Congress was trying to solve on behalf of "everyperson." In contrast, the state governments told stories of federalism and states' rights, which blunted the federal government's story by portraying states as the proper protectors of "everyperson."<sup>180</sup>

In his study, Chestek found patterns and relationships in the litigants' narrative choices and showed how narratives interact depending the category of litigants. He engaged extensively with a small number of subjects based on careful reading and coding of the briefs. His case study developed an in-depth understanding of how lawyers use competing narratives in their briefs.

## 2. *Grounded Theory*

In a grounded theory study, the researcher seeks a general theory of a process or interaction "grounded" in the views of the

---

177. *Id.* at 75–76.

178. Kenneth D. Chestek, *Competing Stories: A Case Study of the Role of Narrative Reasoning in Judicial Decisions*, 9 *LEGAL COMM. & RHETORIC: JALWD* 99 (2012).

179. *Id.* at 109–20.

180. *Id.* at 123–26.

participants. Grounded theory is an effective strategy “when a theory is not yet available to explain a process.”<sup>181</sup>

Grounded theory may involve more subjects than a case study and multiple visits to the field. Rather than choose participants randomly, as in a quantitative study, researchers select the participants most likely to help develop the theory. Interviews are the most common data collection technique in grounded theory.<sup>182</sup>

Creswell describes a “zigzag” data collection process in which the researchers gather information from the participants, analyze the data, then return to gather more information from the participants, and so on. Researchers decide how long to continue data collection by evaluating whether the “categories of information [have] become saturated and whether the theory is elaborated in all of its complexity.”<sup>183</sup> By continually revisiting and refining the data, researchers develop a theory with “specific components: a central phenomenon, causal conditions, strategies, conditions and context, and consequences.”<sup>184</sup>

Erika Abner and Shelley Kierstead recently studied experienced lawyers’ perspectives on novices’ legal writing.<sup>185</sup> They conducted three focus groups made up of fifteen senior Canadian lawyers and one judge. They asked the groups to (1) describe problems with new attorneys’ writing, and (2) assess a sample letter drafted by a relatively new attorney. The researchers then coded the focus group transcripts and identified five broad topics: product, process, identity, teaching and learning, and speculation about why shortcomings exist in novice writing.<sup>186</sup>

The researchers then summarized the themes in the three groups’ comments about each topic. Product comments included problems with grammar, conciseness, organization, analysis, authorities, and attention to facts and audience. Process comments included problems with revision but recognized that budget limitations can constrain the lawyer’s process. Identity comments noted that writing reflects one’s professional self-image. Teaching and learning comments emphasized the value of one-on-one train-

---

181. CRESWELL, *QUALITATIVE INQUIRY*, *supra* note 175, at 62–63, 66.

182. *Id.* at 64–66.

183. *Id.* at 64.

184. *Id.* at 68.

185. Erika Abner & Shelley Kierstead, *A Preliminary Exploration of the Elements of Expert Performance in Legal Writing*, 16 *LEGAL WRITING* 363 (2010).

186. *Id.* at 376–78.

ing and the decline of mentoring. Finally, the experts speculated that novice legal writers have poor reading habits and underestimate the importance of writing.<sup>187</sup> Abner and Kierstead's study illustrates how a grounded theory study can develop themes that inform future work on legal writing expertise.<sup>188</sup>

### 3. *Phenomenological Studies*

Phenomenological research attempts to capture the essence of how the participants experience a particular phenomenon. A phenomenon may be a relatively abstract concept, like insomnia or exclusion, or a more concrete experience, such as undergoing surgery. The researcher engages with a small number of subjects who have lived through the experience in question. The researcher relies primarily on open-ended interviews supplemented by observations and document review and searches for patterns and relationships of meaning in the data that capture the essence of the subjects' experience. A phenomenological study is well-suited for deciding what policies or procedures to develop or for deepening the understanding of a phenomenon.<sup>189</sup>

Although it focused on appellate argument rather than brief writing, a 1996 American Bar Association program illustrated some aspects of a phenomenological study. The program, called "Making the Oral Argument: The View from the Inside Out," is an ABA video accompanied by a handbook. The program involves a simulated appellate argument by two expert appellate practitioners and three current or former judges. In the first portion of the tape, the moderator interviewed the lawyers and judges about their thoughts going into oral argument. The tape then presents the oral argument, followed by a panel discussion reflecting on the argument.

While this program was not strictly a phenomenological study, it offers insight into what such a phenomenological study of legal writing might look like. The program studied subjects who lived an experience: an appellate oral argument. The program involved open-ended interviews of the lawyers and judges before the argument, as well as a panel discussion (akin to a focus

---

187. *Id.* at 389–93.

188. *Id.* at 363.

189. CRESWELL, *QUALITATIVE INQUIRY*, *supra* note 175, at 57–62.

group) after the argument. Finally, the program supplemented those interviews with observation (videotaping the oral argument itself) and document review (the briefs that the lawyers drafted).

Future researchers could apply this structure in a different context—for example, a study of a “real-world” writing project by a cohort of new attorneys. Researchers could conduct interviews or focus groups before and after these writing projects and could supplement the themes from the interviews by reviewing the briefs themselves and possibly observing the arguments, if any, in connection with the cases. The findings might reveal themes or strategies that could benefit future practitioners and help law professors better prepare students for their first real-world writing projects. Alternatively, researchers could conduct a comparative study of both novices and expert practitioners to highlight the differences between how novices and experts experience a writing project.<sup>190</sup>

#### 4. *Ethnographic Research*

In contrast to grounded theory research, which may study disparate participants in the same process, ethnography studies an intact cultural group. The goal is to understand the group’s shared values, behaviors, beliefs, or language. Researchers collect data most commonly through immersive observation and interviews. The research occurs primarily in the participants’ context, that is, where they live or work. The ethnographer’s final product describes the group in detail and analyzes “patterns or topics that signif[y] how the cultural group works and lives.”<sup>191</sup>

Because there appear to be no ethnographic studies of legal writing, an example from educational research helps illustrate ethnographic research. Ann Marie McKee conducted an ethnography of a high school community’s inclusion of a severely disabled student.<sup>192</sup> The student’s disabilities included partial paralysis, visual impairment, and moderate cognitive

---

190. For an example of a phenomenological study in education, see Jennie C. DeGagne & Kelley J. Walters, *The Lived Experience of Online Educators: Hermeneutic Phenomenology*, 6 J. ONLINE LEARNING & TEACHING, June 2010, at 357, 357.

191. CRESWELL, QUALITATIVE INQUIRY, *supra* note 175, at 68–72.

192. Ann Marie McKee, *A Story of High School Inclusion: An Ethnographic Case Study* (unpublished Ph.D. dissertation, Univ. of Iowa 2011), available at <http://ir.uiowa.edu/etd/1247>.

impairment.<sup>193</sup> The student, however, was not the focus of the study.<sup>194</sup> Instead, the researcher focused on the perspective and experience of the parents, administrators, general education teachers, and special education teachers.<sup>195</sup> McKee studied the attitudes of the parents and school personnel toward inclusion, whether the accounts of the participants aligned with or were in tension with one another, and whether that alignment or tension affected the process of inclusion.<sup>196</sup> Over the student's freshman and sophomore years, McKee interviewed 17 parents and professionals involved in the student's inclusion, spent five days observing the student's classes and school activities, and reviewed documents in the school's files.<sup>197</sup>

Among McKee's expansive findings were the following two. First, McKee found that, although the parents and educators were often surprised at how much the student was able to retain, no one—neither parents nor the educators—ever suggested including academic goals in the student's Individualized Education Program.<sup>198</sup> McKee recommended that inclusion efforts must adopt a “presumption of competence” perspective, which in this case would have led to the inclusion of explicit academic goals in the student's Individualized Education Program.<sup>199</sup> Second, McKee noted that, although the participants agreed that inclusion was a laudable goal, everyone except the parents qualified that agreement with concerns about the “disconnect between the theory of inclusion and the practice.”<sup>200</sup> This philosophical difference often led to tension and communication breakdowns between the parents and the administrators and educators.<sup>201</sup> McKee proposed “more open and direct communication” to mitigate the tension and conflict.<sup>202</sup>

McKee's comprehensive study illustrates methods that could translate to the study of legal writers and legal readers. An ethnography studies shared behaviors or values of an intact

---

193. *Id.* at 109.

194. *See id.* at 107.

195. *Id.* at 103.

196. *Id.* at 105–06.

197. *Id.* at 10–11, 115, 122.

198. *Id.* at 201.

199. *Id.* at 203–04.

200. *Id.* at 192.

201. *Id.* at 192, 197.

202. *Id.* at 2.



cultural unit, and the legal profession contains many discrete cultural units for study. Researchers might study how a particular community of judges engages with the briefs they read. It might be especially interesting to the study of that engagement in specialized courts such as juvenile, family law, or drug courts, to see whether the unique demands of those courts affect how the judges engage with the represented and pro se filings. Similarly, researchers could study specific communities of legal writers to see how the demands of their practice areas shape their approach to writing. These researchers might explore the writing practices of local communities of public defenders or prosecutors, large-firm civil litigators, or networks of sole practitioners. Although comparing multiple groups in a single study might be overwhelming, the legal writing field would benefit from multiple researchers studying different groups and learning about the common and disparate beliefs and behaviors in different groups.

##### 5. *Narrative Research*

Narrative research captures the detailed life stories of one or a small number of individuals.<sup>203</sup> Narrative studies may involve a biographical study of one individual or an oral history gathering multiple people's accounts of specific events or time periods.<sup>204</sup> The researcher collects data from multiple sources, such as interviews, observation, and documents. The researcher then reorganizes the various stories into a comprehensive framework. The researcher's retelling may provide "a causal link among ideas" and may highlight themes that arise from the story.<sup>205</sup>

The following primary education study illustrates narrative research. Sonia Houle studied the experiences of a delayed reader, a first-grade boy.<sup>206</sup> Over the boy's first-grade and second-grade years, Houle gathered data from the boy and his parents, teachers, and classmates.<sup>207</sup> She used multiple data sources, including observations, interviews, and analysis of documents

---

203. CRESWELL, *QUALITATIVE INQUIRY*, *supra* note 175, at 54.

204. *Id.* at 55.

205. *Id.* at 56.

206. Sonia T. Houle, *Not Making the Grade: A Narrative Inquiry into Timmy's Experiences with the Mandated Curriculum*, 16 *IN EDUC.*, Autumn 2010, at 30.

207. *Id.* at 33–34.

such as photos, drawings, and schoolwork.<sup>208</sup> The data illustrated the effects of a mandated curriculum on a struggling reader. The mandated curriculum shaped how the teacher and parents unconsciously co-composed a “lived curriculum” with the boy and constructed the boy as a student who was “not good enough” and needed to be fixed.<sup>209</sup>

Narrative research should translate well to the study of legal writing. Legal writing scholars have already studied the final product of great writers.<sup>210</sup> Researchers could go further and study skilled writers to develop a story of their “life in writing” as they progress from novices to experts. These studies could go beyond the writers and their work product and include the experiences of colleagues, adversaries, and even clients and judges to the extent consistent with privileges and codes of conduct.

#### IV. MIXED METHODS RESEARCH

The sections that follow explain the methodology of mixed methods research and illustrate several different types of mixed methods research.

##### A. Basic Methodology of Mixed Methods Research

As its name suggests, mixed methods research combines both quantitative and qualitative analysis. Mixed methods research has become increasingly popular over the past few decades,<sup>211</sup> and 2014 saw the first conference of the Mixed Methods International Research Association.<sup>212</sup> By using both types of approaches, researchers can conduct studies that are stronger than purely quantitative or purely qualitative studies.<sup>213</sup> However, mixed methods research can be more challenging than purely qualitative or quantitative research because of “the need for

---

208. *Id.*

209. *Id.* at 37.

210. *See, e.g.* LINDA H. EDWARDS, READINGS IN PERSUASION: BRIEFS THAT CHANGED THE WORLD (2012); NOAH A. MESSING, THE ART OF ADVOCACY: BRIEFS, MOTIONS, AND WRITING STRATEGIES OF AMERICA’S BEST LAWYERS (2013).

211. CRESWELL, RESEARCH DESIGN, *supra* note 5, at 14–15.

212. Mixed Methods Int’l Research Ass’n, MMRIA Inaugural Conference 2014 Highlights, MMRIA, <http://mmira.wildapricot.org/page-1665517> (last visited June 15, 2015).

213. CRESWELL, RESEARCH DESIGN, *supra* note 5, at 4, 218.

extensive data collection, the time-intensive nature of analyzing both qualitative and quantitative data, and the requirement for the researcher to be familiar with both quantitative and qualitative forms of research.”<sup>214</sup>

## B. Mixed Methods Strategies of Investigation

Creswell identifies three basic mixed methods strategies. The first-*convergent parallel* design-involves a single data collection phase covering both quantitative and qualitative data.<sup>215</sup> The second and third-*explanatory sequential* and *exploratory sequential*-involve two separate data collection phases, one quantitative and one qualitative.<sup>216</sup> The next three subsections explain and illustrate these basic mixed methods strategies.

### 1. Explanatory Sequential Design

An explanatory sequential study begins with a quantitative phase and ends with a qualitative phase that explores some aspect of the quantitative findings.<sup>217</sup> For example, the quantitative phase might find an unexpected correlation. The subsequent, qualitative phase might involve interviews of a subset of respondents to generate a theory to explain the correlation. This strategy typically appeals to researchers with stronger quantitative backgrounds.<sup>218</sup> An explanatory sequential study can be easier to execute than a convergent parallel study because one data collection phase builds on the other.<sup>219</sup> However, an explanatory sequential study can take longer to implement because of the successive data collection phases .

Scott Moss conducted an explanatory sequential study of whether low-quality plaintiffs’ briefs were more likely to lose on

---

214. *Id.* at 218–19.

215. *Id.* at 227–28.

216. *Id.* at 15–16. Creswell describes several advanced mixed methods designs that are beyond the scope of this Article. A *transformative design* incorporates a convergent or sequential design “within a social justice framework to help a marginalized group.” *Id.* at 228. A *multiphase design* includes several mixed methods projects “in a longitudinal study with a focus on a common objective for the multiple projects.” *Id.*

217. *Id.* at 224.

218. *Id.*

219. *Id.* at 225.

summary judgment in employment discrimination cases.<sup>220</sup> He drew a random sample from district courts in the Second and Seventh Circuits. The sample included only summary judgment briefs from cases in which the employer relied on the “same actor” defense, an issue on which there was an intra-circuit split in each circuit. Given the intra-circuit split, plaintiff’s counsel could not write an effective brief without citing the authority within his circuit rejecting the defense. Moss found that 73% of plaintiffs’ briefs failed to cite the pro-plaintiff case law on the same actor defense.<sup>221</sup> In addition, Moss found that briefs omitting the pro-plaintiff authority lost at roughly double the rate (86%) of plaintiffs whose counsel cited that authority (42%), a statistically significant difference ( $p < .0001$ ).<sup>222</sup> After tabulating this quantitative data, Moss collected qualitative data about the quality of the “bad” briefs.<sup>223</sup> He reported and offered examples of incompetent writing laden with grammatical flaws, unnecessary and fatal concessions, inadequate and outdated legal research, and outright procedural default.<sup>224</sup> Based on these findings, Moss engaged in a comprehensive discussion of potential causes and solutions to the problem.<sup>225</sup>

## 2. *Exploratory Sequential Design*

An exploratory sequential study begins with a qualitative phase and ends with a quantitative phase informed by the qualitative findings.<sup>226</sup> Researchers may use the initial qualitative phase to identify variables to study in the quantitative phase, to identify appropriate instruments to use in the quantitative phase, or even to develop a new instrument for the quantitative phase.<sup>227</sup> For example, Moss could have begun his

---

220. Scott A. Moss, *Bad Briefs, Bad Law, Bad Markets: Documenting the Poor Quality of Plaintiffs’ Briefs, Its Impact on the Law, and the Market Failure It Reflects*, 63 EMORY L.J. 59 (2013).

221. *Id.* at 81.

222. *Id.* at 90.

223. *Id.* at 81. Moss classified briefs as “bad” based on their failure to include case law or argument against the same-actor defense. *Id.* at 80. His additional examination of research, writing, and strategic blunders offered the reader a qualitative perspective on the “badness” of these briefs. *Id.* at 80–81.

224. *Id.* at 82–90.

225. *Id.* at 94–123.

226. CRESWELL, RESEARCH DESIGN, *supra* note 5, at 225–26.

227. *Id.* at 16.

study with a qualitative phase to develop a theory of what makes briefs likely to lose and then used that theory to decide what quantitative data to collect and analyze. Like the explanatory sequential strategy, the two-stage data collection may make the exploratory sequential strategy easier to execute than the convergent parallel strategy.<sup>228</sup> Yet the two-stage data collection model take longer.

### 3. *Convergent Parallel Design*

Researchers in a convergent parallel study collect qualitative and quantitative data simultaneously.<sup>229</sup> Researchers must collect both types of data “using the *same or parallel variables, constructs, or concepts*.”<sup>230</sup> In theory, these two different types of data should yield similar results.<sup>231</sup> The researchers then analyze the quantitative and qualitative data in search of convergences or differences.<sup>232</sup>

Judith Fisher’s study of issue statements in state supreme court briefs has elements of a convergent parallel design.<sup>233</sup> She sampled briefs in 300 cases filed in six state supreme courts. She then examined eight different characteristics of the issue statements in those briefs: words per question, issues per question, clarity, sentence structure, opening words, party designations, inclusion of facts, and persuasive techniques. For some characteristics, Fisher tabulated quantitative data across the entire sample and broke down the tabulations by state. For clarity, Fisher described examples of both clear and unclear questions.<sup>234</sup> And for persuasiveness, Fisher took both approaches; she calculated the percentage of questions phrased so that a “yes” answer favored the client, and she examined more subjective word choice decisions. By combining numerical measures with carefully selected examples, Fisher used triangulation to offer a fuller

---

228. *Id.* at 225.

229. *Id.* at 219.

230. *Id.* at 222 (emphasis in original).

231. *Id.* at 219.

232. *Id.* at 223.

233. Judith D. Fisher, *Got Issues? An Empirical Study About Framing Them*, 6 J. ALWD 1, 4–5 (2009).

234. *Id.* at 9–10.

picture of the questions presented than a purely quantitative or qualitative study could have produced.

Chestek's study of whether narrative elements make briefs appear more persuasive offers another example of a convergent parallel.<sup>235</sup> Chestek surveyed judges, clerks, and practitioners about which briefs they found more persuasive—a “story” brief that included a strong narrative element or a “logos” brief that did not.<sup>236</sup> The survey also included an open-ended “comment” field.<sup>237</sup> Chestek used the open-ended comments to understand whether the respondents' stated preferences were really based on the presence or absence of a “story,” as opposed to some other aspect of the briefs.<sup>238</sup> Those who preferred the “story” brief often commented on the context or personalization that the brief provided.<sup>239</sup> In contrast, those preferring the “logos” brief often commented that the “story” brief included irrelevant or unnecessary facts, whereas the “logos” brief properly focused on the legal issues.<sup>240</sup> Thus, the qualitative data told a similar story as the quantitative data told about the preference for narrative in persuasive writing.

## V. PRACTICAL ADVICE FOR PLANNING AN EMPIRICAL STUDY

The sections that follow offer some practical advice for new empirical researchers concerning potential empirical research questions, data collection issues, the role of institutional review boards and methodologists, and writing up a study.

### A. Potential Research Questions

Your choice of research question will be determined largely by your interests and creativity, as well as the resources available to you. Nevertheless, it may be helpful to consider several different types of questions that empirical research might answer.

One set of questions could explore the relationship between legal writing practices and case outcomes. Several different

---

235. Chestek, *supra* note 135.

236. *Id.* at 8.

237. *Id.* at 22.

238. *Id.* at 22–24.

239. *Id.*

240. *Id.*

quantitative techniques might help answer these types of questions. Researchers might conduct correlational studies of existing cases, as Long and Christenson did in their study of readability and success on appeal<sup>241</sup> or as Moss did in his mixed-methods study of employment discrimination summary judgment motions.<sup>242</sup> Or they might conduct an experiment asking participants to simulated cases, as in the proposed modification of Chestek's study of preferences for narrative elements in briefs.<sup>243</sup> Alternatively, researchers could conduct qualitative studies of decision-makers believe that legal writing influences their decisions. The most likely research designs would seem to be case studies and phenomenological studies. However, any study in this area must be mindful that the subjects' stated preferences for a given brief may not accurately predict whether that brief would persuade them.

A second set of questions might involve the relationship between legal writing instruction and the practice of legal writing. Has the gradual inclusion of legal writing instruction into the mainstream of legal education changed how practicing lawyers write? Ian Gallacher considered this question in his study of readability scores of briefs filed in New York's highest court from 1969 to 2008.<sup>244</sup> Gallacher's study "was designed to reveal if the effects of systematic legal-writing instruction in law schools could be seen in documents written by lawyers."<sup>245</sup> Gallacher's findings suggest "that the trend is actually moving away from plainer writing, even at a time when legal-writing teachers' efforts should be producing the opposite effect."<sup>246</sup> Similarly, in two separate studies of the ways that appellate practitioners draft questions presented, Brady Coleman and Judith Fisher compared the "best

---

241. Long & Christensen, *supra* note 126.

242. Moss, *supra* note 220.

243. See Chestek, *supra* note 146.

244. Ian Gallacher, "When Numbers Get Serious": A Study of Plain English Usage in Briefs Filed Before the New York Court of Appeals, 46 SUFFOLK U. L. REV. 451, 462-63 (2013). Gallacher collected eight briefs from each year and sampled three pages from the argument section to generate the readability score, as calculated by Microsoft Word 2003, as well as words per sentence, sentences per paragraph, and incidence of passive voice. Gallacher averaged the scores of each year's briefs to yield an average score for the year. *Id.* at 463-64. Gallacher's discussion of his findings addresses a variety of potential reasons why readability might have increased over those forty years, many of which provide interesting avenues for future research. *Id.* at 491 n.98.

245. *Id.* at 460.

246. *Id.* at 457.

practices” urged by professors and expert practitioners with the briefs that practitioners actually filed.<sup>247</sup>

Researchers could also use qualitative research designs to study the whether legal writing pedagogy influences practice. Researchers might use case studies to explore how practitioners connect their legal writing education to their law practice or to explore how legal writing practices have evolved over time. They might conduct narrative research studying a few practitioners or a few legal organizations over many years to see how their writing practices have evolved.

A third set of questions may involve whether writing practices vary based on the setting. For example, researchers might explore whether lawyers write differently based on the type of case, the size of law firm, or the party represented. Researchers could identify the standard “playbook” that lawyers in a given jurisdiction use in particular types of cases. For example, one might sample briefs in employment discrimination cases to reveal patterns in how the plaintiffs’ and defendants’ bar handle those cases. Surveys, correlational studies, and experiments could all shed light on these questions, as could case studies and ethnographies.

## B. Data Collection Issues

After choosing a topic to investigate, the researcher must next consider what data are available about the topic. When studying legal writing, these data likely fall into two broad categories: documents (the legal writing itself), and the people who write and read those documents.

When gathering data from legal documents, researchers may take advantage of increasingly effective text analysis tools. Michael Evans and his colleagues provided an overview of text analysis (a/k/a “text classification”) tools in their article on using automated text analysis tools to study Supreme Court briefs.<sup>248</sup> The goal of text analysis, they explained, “is to develop automatic methods for labeling previously unseen documents according to

---

247. Coleman et al., *supra* note 104, at 327; Fisher, *supra* note 233, at 4–5.

248. Michael Evans et al., *Recounting the Courts? Applying Automated Content Analysis to Enhance Empirical Legal Research*, 4 J. EMPIRICAL LEGAL STUD. 1007 (2007).



some predefined coding scheme, where the labels are drawn from a finite set of alternatives.”<sup>249</sup>

At one end of the spectrum, text-analysis software tools can be quite simple. For example, Allison Martin used Wordle’s word-counting function to study the themes in a set of trial-level briefs in litigation challenging the Affordable Care Act.<sup>250</sup> Wordle is a text visualization tool that counts the number of times words appear in a given text and displays “word clouds.” The most frequently used words appear in the largest type, while less frequently used words appear in increasing smaller type.<sup>251</sup> Martin found that the themes that Wordle’s word clouds revealed were consistent with the themes identified in Kenneth Chestek’s 2012 study of the same set of briefs.<sup>252</sup>

At the other end of the spectrum, text analysis software can be highly complex and customized. For example, in Coleman and Phung’s study of plain language usage in 9,000 Supreme Court briefs from 1969 to 2004, they wrote eight PERL<sup>253</sup> software programs to clean up, sort, and analyze the briefs.<sup>254</sup> Their customized text analysis tools detected not only four different readability measures but also five categories of terms that indicate complexity – the use of “stuffy” terms with plainer counterparts, compound constructions, redundant legal phrases, “Lawyerisms,” and “Latinisms.”<sup>255</sup>

---

249. *Id.* at 1010.

250. Allison D. Martin, *A Picture Is Worth a Thousand Words: How Wordle™ Can Help Legal Writers*, 9 LEG. COMM. & RHETORIC: JALWD 139 (2012).

251. *Id.* at 140.

252. *Id.* at 142–46 (citing Chestek, *supra* note 178). Another relatively simple set of tools is Microsoft Word’s set of readability scores and word counts, used in Long & Christensen, *supra* note 126, and Gallacher, *supra* note 237.

253. “Perl is a general-purpose programming language originally developed for text manipulation and now used for a wide range of tasks including system administration, web development, network programming, GUI development, and more.” Perl 5 Version 20.1 Documentation, *What Is Perl*, <http://perldoc.perl.org/perlintro.html> (last visited June 15, 2015).

254. Brady Coleman & Ouy Phung, *The Language of Supreme Court Briefs: A Large-Scale Quantitative Investigation*, 11 J. APP. PRAC. & PROCESS 75, 76, 78 (2010). Coleman & Phung developed an automated process to replace citations with the term “scite,” which eliminated uncertainty about how their text analysis tools would process the various legal citations. *Id.* at 81.

255. *Id.* at 82–85. For a sampling of text analysis tools, see the Stanford Natural Language Processing Group’s online collection of resources, available at <http://nlp.stanford.edu/software/index.shtml>.

Alternatively, researchers may rely on manual document review and coding. Although manual review and coding may be time consuming and impose practical limitations on how many briefs one can study, it allows researchers to investigate document features of interest for which no automated tests exist. For example, Michael Murray used manual review and coding to study the use of parentheticals in 200 federal appellate court briefs and 107 Supreme Court opinions.<sup>256</sup>

When gathering data from people, researchers have multiple options. They may rely on surveys, whether conducted in-person, by mail, or by email. They may also choose to conduct individual interviews or focus groups. And they may conduct assessments of individuals' knowledge, skills, beliefs, or practices.

Researchers designing a document-based study should ask some important questions at the outset. Who has custody of the documents? Can researchers access them? Are the documents indexed or classified in any way? Can researchers obtain the documents in digital form and, if so, in what format or formats? If the documents are images, are they of sufficiently high quality to permit reliable optical character recognition? Will researchers need access to the complete set of documents to permit random sampling, or will a pre-selected subset suffice?

Researchers must also consider what it will cost to access the documents. Most legal academics have free access to briefs, pleadings, and motions on WestlawNext or Lexis Advance, but those services may not provide comprehensive databases containing all filings for a given jurisdiction. For federal court filings, PACER offers a complete record, but can be prohibitively expensive.<sup>257</sup>

After selecting a research strategy, the researcher will have to decide who will collect the data. That will depend largely on the skills and experience required by the data collection method. For many quantitative data collection tasks, student research assistants will likely be able to collect the data, under the researcher's careful and frequent supervision. Researchers should spot-check the students' findings early and often in order to catch and resolve potential problems before wasting precious time.

---

256. Murray, *supra* note 110, at 231–33.

257. See PACER Electronic Public Access Fee Schedule, [https://www.pacer.gov/documents/epa\\_feesched.pdf](https://www.pacer.gov/documents/epa_feesched.pdf) (last visited June 15, 2015) (stating fee of \$0.10 per page up to 30 pages per document).

For most quantitative collection, the real skill comes in designing the data collection instrument. The instrument may be a survey, which requires careful consideration and testing to avoid wasting time and resources.<sup>258</sup> The instrument may be a protocol that walks researchers through the relatively mechanical process of selecting documents and measuring the variables of interest. Or the instrument may be a text-analysis software tool that requires some expertise to design or customize.

When collecting qualitative data from documents, the proper research assistant will depend on the nature of the coding task. For fairly routine tasks, student research assistants may be qualified based on several years of education in legal writing and analysis. However, for more sophisticated data gathering that require an experienced legal reader or writer, a law student may be insufficient, and the researcher may want to review all of the documents personally.

Similarly, interviews or focus groups may not be suited to a law student's skill set, except perhaps for law students with a strong background in qualitative research. Even those students, however, may lack the necessary subject matter expertise to conduct effective interviews. If the qualitative data gathering task is too big to accomplish individually, consider partnering with one or more legal writing colleagues. In addition to dividing the data collection task, researchers will benefit from adding multiple perspectives to the study.

### C. Institutional Review Boards

If you are engaged in "human subject research," you may be subject to rules established by your institution's Institutional Review Board ("IRB"). Technically, the federal IRB regulations apply only to federally funded research.<sup>259</sup> Institutions receiving federal funds must provide assurances that they will protect human subjects in all of their research, whether or not it is federally funded. In practice, therefore, many IRBs apply the federal regulatory approach to all of their research.<sup>260</sup>

---

258. For a comprehensive discussion of survey methodology, see NEUMAN, *supra* note 24, at 263–307.

259. David A. Hyman, *Institutional Review Boards: Is This the Least Worst We Can Do?*, 101 NW U. L. REV. 749, 752 (2007).

260. *Id.*

If your study will involve only publicly available documents, you likely do not fall within the definition of “human subject” research. However, if your study will gather data from people or will gather data about people from private documents, you could be conducting human subject research. The regulations define a human subject as “a living individual about whom an investigator (whether professional or student) conducting research obtains (1) [d]ata through intervention or interaction with the individual, or (2) [i]dentifiable private information.”<sup>261</sup>

Even if your research falls within an exemption under your institution’s IRB rules, there is almost certainly a process to secure your IRB’s determination of exemption. Therefore, if any chance exists that your study involves human subject research, you should consult your IRB early and often. You may be well served to seek guidance from another researcher at your institution with substantial experience dealing with IRB policies and practices at your institution.

#### D. The Role of a Methodologist

The ideal methodologist is someone experienced with empirical research design. Even if your project is quantitative, the methodologist need not be a statistician. Academics in many other fields rely on quantitative and qualitative research methods. And some of those fields lend themselves to interdisciplinary studies involving legal writing, such as psychology, sociology, criminal justice, and political science.

Unless you have experience conducting empirical research, consider involving a methodologist as early as possible. Early collaboration will not only help avoid wasting time and resources but may also suggest research designs that would not have occurred to you. Collaboration may be as complete as co-authorship of the study or as minimal as asking for guidance in designing the study and analyzing your data.

As you consider reaching out to a methodologist, you might wonder why someone would want to help with your project. A methodologist might have several reasons to help beyond mere

---

261. 45 C.F.R. § 46.102(f) (2014). Any empirical study would meet the regulatory definition of “research” as “a systematic investigation, including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge.” 45 C.F.R. § 46.102(d).

altruism. First, especially for untenured faculty, a publishing opportunity might be attractive. Moreover, many universities are pushing interdisciplinary collaboration, and your institution may support or reward collaborations with other schools. Finally, an interdisciplinary study involving legal writing and the legal system may simply be interesting to faculty in other disciplines.

#### E. Writing up the Study

The format for writing up a study varies depending on the research methods and investigative strategies used. One strategy for selecting a format is to find published studies with similar research strategies and use those as a model. Many of the empirical studies cited above have appeared in *Legal Communication and Rhetoric: JALWD*, or in *Legal Writing: The Journal of the Legal Writing Institute*. Another potential source of examples is the *Journal of Empirical Legal Studies*. For excellent practical advice on writing up a variety of studies, see John W. Creswell, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (4th ed. 2014).

### VI. CONCLUSION

Empirical research methods hold great promise for the study of legal writing. In recent years, legal writing scholars have done outstanding theoretical work drawing on other disciplines. Empirical research allows legal writing scholars to test how those theories work in the real world, to study the influence of legal writing pedagogy on law practice, and to deepen the field's understanding of legal writers and legal readers.

The examples described in this Article paint a picture of a vibrant and growing field of research. My hope is that this Article will help spur new entrants into this field, as well as highlight how empirical research can contribute to legal writing scholarship.

2015      *Using Empirical Methods to Study Legal Writing*      185

## Appendix

### A. General Resources on Empirical Research Design

John W. Creswell, *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research* (4th ed. 2011).

John W. Creswell, *Qualitative Inquiry & Research Design: Choosing Among Five Approaches* (3d ed. 2012).

John W. Creswell, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (4th ed. 2014).

Daniel Muijs, *Doing Quantitative Research in Education with SPSS* (2d ed. 2011).

W. Lawrence Neuman, *Social Research Methods: Qualitative and Quantitative Approaches* (7th ed. 2009).

R. Murray Thomas, *Blending Qualitative and Quantitative Research Methods in Theses and Dissertations* (2003).

*The Oxford Handbook of Empirical Legal Research* (Cane & Kritzer eds., 2010).

John W. Creswell, *Mixed Methods Research, Videos*, <http://johnwcreswell.com/videos> (video introducing mixed methods research)

### B. Empirical Studies of Legal Writing

Erika Abner & Shelley Kierstead, *A Preliminary Exploration of the Elements of Expert Performance in Legal Writing*, 16 *Legal Writing* 363 (2010).

Robert W. Benson & Joan B. Kessler, *Legalese v. Plain English: An Empirical Study of Persuasion and Credibility in Appellate Brief Writing*, 20 *Loy. L.A. L. Rev.* 301 (1987).

- Charles A. Bird & Webster Burke Kinnaird, *Objective Analysis of Advocacy Preferences and Prevalent Mythologies in One California Appellate Court*, 4 J. App. Prac. & Process 141 (2002).
- Robin A. Boyle & Joanne Ingham, *Suggestions on How to Conduct Empirical Research: A Behind-the-Scenes View*, 15 Persps. 176 (2007).
- Kenneth D. Chestek, *Competing Stories: A Case Study of the Role of Narrative Reasoning in Judicial Decisions*, 9 Legal Comm. & Rhetoric: JALWD 99 (2012).
- Kenneth D. Chestek, *Judging by the Numbers: An Empirical Study of the Power of Story*, 7 J. ALWD 1 (2010).
- Brady S. Coleman et al., *Grammatical and Structural Choices in Issue Framing: A Quantitative Analysis of "Questions Presented" from Half a Century of Supreme Court Briefs*, 29 Am. J. Trial Advoc. 327 (2005).
- Brady Coleman & Quy Phung, *The Language of Supreme Court Briefs: A Large-Scale Quantitative Investigation*, 11 J. App. Prac. & Process 75 (2010).
- Pamela Corley, et al., *The Influence of Amicus Curiae Briefs on U.S. Supreme Court Opinion Content* (unpublished manuscript delivered at the 109th Annual Meeting of the American Political Science Ass'n, Chicago, IL, Aug. 31 2013), available at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2300505](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2300505)).
- Linda H. Edwards, *Readings in Persuasion: Briefs That Changed the World* (2012).
- Michael Evans et al., *Recounting the Courts? Applying Automated Content Analysis to Enhance Empirical Legal Research*, 4 J. Empirical Legal Stud. 1007 (2007).
- Judith D. Fischer, *Got Issues? An Empirical Study About Framing Them*, 6 J. ALWD 1 (2009).

2015      *Using Empirical Methods to Study Legal Writing*      187

Sean Flammer, *Persuading Judges: An Empirical Analysis of Writing Style, Persuasion, and the Use of Plain English*, 16 *Legal Writing* 183 (2010).

Ian Gallacher, “When Numbers Get Serious”: *A Study of Plain English Usage in Briefs Filed Before the New York Court of Appeals*, 46 *Suffolk U. L. Rev.* 451 (2013).

Joseph Kimble, *The Straight Skinny on Better Judicial Opinions (Part 1)*, in *Lifting the Fog of Legalese: Essays on Plain Language* 15–35 (2006).

Joseph Kimble, *The Straight Skinny on Better Judicial Opinions (Part 2)*, in *Lifting the Fog of Legalese: Essays on Plain Language* 89–104 (2006).

Joseph Kimble, *Strike Three for Legalese*, in *Lifting the Fog of Legalese: Essays on Plain Language* 3–13 (2006).

Lance N. Long & William F. Christensen, *Clearly, Using Intensifiers Is Very Bad—Or Is It?*, 45 *Idaho L. Rev.* 171 (2008).

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

Lance N. Long & William F. Christensen, *When Judges (Subconsciously) Attack: The Theory of Argumentative Threat and the Supreme Court*, 91 *Or. L. Rev.* 933 (2013).

Allison D. Martin, *A Picture Is Worth a Thousand Words: How Wordle™ Can Help Legal Writers*, 9 *Leg. Comm. & Rhetoric: JALWD* 139 (2012).

Noah A. Messing, *The Art of Advocacy: Briefs, Motions, and Writing Strategies of America’s Best Lawyers* (2013).

Scott A. Moss, *Bad Briefs, Bad Law, Bad Markets: Documenting the Poor Quality of Plaintiffs’ Briefs, Its Impact on the Law, and the Market Failure It Reflects*, 63 *Emory L.J.* 59 (2013).



- Michael D. Murray, *The Promise of Parentheticals: An Empirical Study of the Use of Parentheticals in Federal Appellate Briefs*, 10 *Legal Comm. & Rhetoric: J. ALWD* 229 (2013).
- Michael Rustad & Thomas Koenig, *The Supreme Court and Junk Social Science: Selective Distortion in Amicus Briefs*, 72 *N.C. L. Rev.* 91 (1993).
- Stacy Rogers Sharp, *Crafting Responses to Counterarguments: Learning from the Swing Vote Cases*, 10 *Legal Comm. & Rhetoric: J. ALWD* 201 (2013).
- Gregory C. Sisk & Michael Heise, *"Too Many Notes": An Empirical Study of Advocacy in Federal Appeals*, 12 *J. Empirical Legal Stud.* (forthcoming 2015).