

**GATEKEEPER WITH A GAVEL: A SURVEY  
EVALUATING JUDICIAL MANAGEMENT  
OF CHALLENGES TO EXPERT  
RELIABILITY AND THEIR RELATIONSHIP  
TO SUMMARY JUDGMENT**

*Andrew Jurs\**

INTRODUCTION.....	326
I. PRIOR STUDIES ON JUDICIAL HANDLING OF RELIABILITY MOTIONS .....	328
<i>A. Reliability Motions – Empirical Studies.....</i>	<i>329</i>
1. Frequency of <i>Daubert</i> Gatekeeping Questions in Court .....	330
2. Procedural Methodology of Handling Gatekeeping.....	332
<i>B. Substantive Factors Used to Evaluate Reliability.....</i>	<i>333</i>
II. RELIABILITY MOTIONS AND SUMMARY JUDGMENT: COMMENTARY & PRIOR STUDIES .....	335
III. A NEW SURVEY ON JUDICIAL HANDLING OF GATEKEEPING .....	339
<i>A. Study Methodology.....</i>	<i>339</i>
<i>B. Analysis of Responses.....</i>	<i>343</i>
<i>C. Judicial Handling of Reliability Motions.....</i>	<i>344</i>
<i>D. Judicial Handling of Summary Judgment .....</i>	<i>352</i>
<i>E. Comparing Reliability Motion Data to Summary        Judgment Data.....</i>	<i>357</i>
<i>F. Comparing Judges Based on Home-State        Admissibility Standard.....</i>	<i>364</i>
IV. DISCUSSION AND SUGGESTIONS FOR FURTHER RESEARCH.....	367

---

\* Associate Professor of Law, Drake University Law School; J.D., University of California, Berkeley, School of Law; B.A., Stanford University. The author wishes to thank Kevin Saunders and Allan Vestal for their comments on an earlier draft of this work. I also wish to thank Krissa LeLaCheur Mason for her outstanding research assistance for this project. Thanks also to Katie, Clara, and Milo.

A. Implications of the Survey Data .....	367
B. Suggestions for Future Research .....	369
CONCLUSION .....	371

## INTRODUCTION

*Daubert's* gatekeeping standard has, since 1993, granted judges extraordinary authority to screen expert evidence in federal court. To screen the evidence, judges have been given a series of substantive factors to consider,<sup>1</sup> and the Supreme Court has suggested several procedural methods to evaluate complex evidence.<sup>2</sup> Beyond that guidance, judges have a wide range of discretion to do their work.<sup>3</sup>

In the decade after *Daubert*, a series of studies inquired into judicial handling of the new gatekeeping authority.<sup>4</sup> The studies evaluated the frequency of motions challenging expert evidence, the procedural methods used to evaluate evidence, and the substantive factors judges used to decide contested expert issues. These studies were extraordinarily important in understanding how the landscape of expert witnesses had changed in the period of 1990 to 2000. But many studies explicitly recognize that additional research is needed,<sup>5</sup> and the age of the studies suggest updating their findings.<sup>6</sup>

---

<sup>1</sup> *Daubert v. Merrell Dow Pharm.*, 509 U.S. 579, 593-95 (1993) (discussing the “Daubert factors” for reliability assessment). For further discussion of these factors, see *infra* note 13 and accompanying text.

<sup>2</sup> *Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 147 (Breyer, J., concurring). For further discussion of these methods, see *infra* note 14 and accompanying text.

<sup>3</sup> In *Joiner*, the Court established the abuse of discretion standard for appellate review of *Daubert* gatekeeping. *Joiner*, 522 U.S. at 139.

<sup>4</sup> See, e.g., LLOYD DIXON & BRIAN GILL, RAND INSTITUTE FOR CIVIL JUSTICE, CHANGES IN THE STANDARDS FOR ADMITTING EXPERT EVIDENCE IN FEDERAL CIVIL CASES SINCE THE *DAUBERT* DECISION (2001); Shirley A. Dobbin et al., *Federal and State Trial Judges on the Proffer and Presentation of Expert Evidence*, 28 JUST. SYS. J. 1 (2007); Jennifer Groscup et al., *The Effects of Daubert on the Admissibility of Expert Testimony in State and Federal Criminal Cases*, 8 PSYCHOL. PUB. POL'Y & L. 339 (2002); Carol Krafka et al., *Judge and Attorney Experiences, Practices, and Concerns Regarding Expert Testimony in Federal Civil Trials*, 8 PSYCHOL. PUB. POL'Y & L. 309 (2002); Sophia Gatowski et al., *Asking the Gatekeepers: A National Survey of Judges on Judging Expert Evidence in a Post-Daubert World*, 25 LAW & HUM. BEHAV. 433 (2001).

<sup>5</sup> See, e.g., Dobbin et al., *supra* note 4, at 14 (“Additional research should build upon these findings to provide further insight into the admissibility process and how

Beyond the descriptive analyses of reliability motions, other commentators have also connected these reliability motions to the current resurgence of summary judgment.<sup>7</sup> Reliability analysis of expert testimony, pursuant to *Daubert* gatekeeping, granted judges a powerful tool for case management. Judges accepting the invitation to manage cases may be one explanation for the increased use of summary judgment in the post-*Daubert* years.<sup>8</sup> Yet for all the commentary, the connection between reliability and summary judgment motions has not been empirically analyzed.

Using a survey methodology, this Study will analyze reliability motions in trial courts by examining the frequency, procedural methodology, and substantive factors used to handle those motions.<sup>9</sup> That portion of the Study will update the prior research in the area, shed light on the judicial handling of reliability motions now that *Daubert* is twenty years old, and provide a comparison to previous data to see if judicial gatekeeping has changed. Beyond that analysis, the Study will

---

judges manage the expert testimony before them, especially now that *Daubert* is such an accepted part of the legal landscape.”); Krafka et al., *supra* note 4, at 330-31 (“Although this work addresses an important set of issues, it describes only the final stages of litigation involving expert evidence. To determine how *Daubert* and its associated cases have affected judicial and attorney practices in the majority of cases that never go to trial, further research is needed.”).

<sup>6</sup> For example, the survey data discussed *infra* Part II.A., was taken in the 1990’s. Dobbins et al., *supra* note 4, at 3 (1999 judicial survey); Krafka et al., *supra* note 4, at 311 (1991 and 1998 judicial surveys).

<sup>7</sup> See, e.g., Lucinda Finley, *Guarding the Gate to the Courthouse: How Trial Judges are Using their Evidentiary Screening Role to Remake Tort Causation Rules*, 49 DEPAUL L. REV. 335, 336 (1999) (“While adopting substantive changes in causation law through the rubric of evidentiary admissibility decisions, judges have also frequently conflated admissibility decisions and sufficiency of evidence decisions.”); Peter H. Schuck, *Judicial Avoidance of Juries in Mass Tort Litigation*, 48 DEPAUL L. REV. 479, 498 (1998) (courts use *Daubert* gatekeeping to exclude types of plaintiffs’ evidence, to avoid trial); Joseph Sanders, *Scientific Validity, Admissibility, and Mass Torts After Daubert*, 78 MINN. L. REV. 1387, 1433 (1994) (“[C]ourts have attempted to achieve with admissibility rulings what they should be achieving with sufficiency rulings.”).

<sup>8</sup> DIXON & GILL, *supra* note 4, at 56-57; Joe Cecil et al., *A Quarter-Century of Summary Judgment Practice in Six Federal District Courts*, 4 J. EMPIRICAL LEG. STUD. 861, 882-83, 896 (2007). *But see* Theodore Eisenberg & Charlotte Lanvers, *Summary Judgment Rates Over Time, Across Case Categories, and Across Districts: An Empirical Study of Three Large Federal Districts*, available at [http://scholarship.law.cornell.edu/lrsp\\_papers/108/](http://scholarship.law.cornell.edu/lrsp_papers/108/). For a more detailed review of empirical data on summary judgment rates, see *infra* Part II.B.

<sup>9</sup> See *infra* Part III.

also offer some preliminary findings on how judges handle summary judgment, and start to measure the connection between reliability screening and summary judgment.<sup>10</sup> That section also begins to assess whether the commentary on the issue can be empirically verified.

To these ends, I will begin in Part II with a brief outline of the prior research in the area of reliability motions and how they have been studied in the period after *Daubert*. The literature review will also include some of the analysis of summary judgment, and also the commentary discussing the connection of reliability to Rule 56 motions.<sup>11</sup> In Part III, I will discuss the survey instrument used to ask judges about their handling of these motions, and also discuss the results of the survey from all judges. I will finish in Part IV with a discussion of the findings of the survey, and then comment on further research suggested by this Study.

By measuring the actual practices of judges handling reliability motions, this survey offers important new data exploring into how judges assess litigants' challenges and decide those motions, and also offers some clues as to whether commentators are right when they suggest gatekeeping has expanded to include inquiry into factual sufficiency.

#### I. PRIOR STUDIES ON JUDICIAL HANDLING OF RELIABILITY MOTIONS

The Supreme Court decision in *Daubert v. Merrell Dow Pharm.* was a watershed moment for expert evidence in federal court. In their decision, the Court mandated a new process for evaluating proposed expert testimony, necessitating relevance and reliability screening prior to admission.<sup>12</sup> The gatekeeping

---

<sup>10</sup> See *infra* Parts II.B (commentary on relationship of reliability to summary judgment) and III.B (survey results). See also Michael D. Green, *The Impact of the Civil Jury on American Tort Law*, 38 PEPP. L. REV. 337, 356 (2011) (stating past handling of motions may be the best method of predicting future handling of motions).

<sup>11</sup> See *infra* Part II.B.

<sup>12</sup> *Daubert v. Merrell Dow Pharm.*, 509 U.S. 579, 589 (1993) (FED. R. EVID. 702 provides the standard for expert admissibility, rather than "general acceptance" standard).

function would permit judges to assess the merits of the science offered in court to ensure it met minimum standards.<sup>13</sup>

To these ends, the Court has suggested both substantive factors to consider in gatekeeping and procedural methods to assist in that decision. To evaluate the reliability of evidence, the Court offered the following substantive factors: technique can and has been tested, peer review and publication, known or potential rate of error, standards for controlling the technique, and general acceptance within the scientific community.<sup>14</sup> In a later case, *General Electric v. Joiner*, Justice Breyer offered thoughts on the procedures that may assist a judge in those findings. Justice Breyer suggested that judges could use special masters, pretrial conferences, questioning from the bench, or independent experts to make findings on reliability; each of these methods is contained within the Rules of Civil Procedure or Rules of Evidence.<sup>15</sup>

Between 1993 and 2002, scholars and researchers repeatedly evaluated the effect of the change in judicial screening mandated by *Daubert*. Researchers measured the changes *Daubert* unleashed in courtroom practice and case management by analyzing the effects on judges handling expert reliability questions.<sup>16</sup> These studies were critically important in measuring the initial effects of *Daubert* and understanding how it changed judicial gatekeeping.

#### A. Reliability Motions – Empirical Studies

To evaluate these studies, I will break them into several different subjects, discussing first the frequency of reliability issues arising, then reviewing the procedures judges use to handle gatekeeping. Finally, I will recap the prior work discussing the substantive factors used by judges to decide reliability.

---

<sup>13</sup> *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 152 (1999) (stating *Daubert* gatekeeping standard exists “to make certain that an expert . . . employs in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field”).

<sup>14</sup> *Daubert*, 509 U.S. at 593-95.

<sup>15</sup> *Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 149 (Breyer, J., concurring).

<sup>16</sup> *See, e.g.*, Dobbin et al., *supra* note 4, at 1-2 (survey intended to empirically measure strategies used by judges in court after *Daubert*); Krafka et al., *supra* note 4, at 309-11 (survey research intended to see effect of *Daubert* on judges and attorneys).

### 1. Frequency of *Daubert* Gatekeeping Questions in Court

One question researchers wanted to measure was how often a reliability issue would arise in court. To measure the frequency, different studies used distinct methodologies.

Lloyd Dixon and Brian Gill measured the frequency of reliability determinations using statistical analysis of cases appearing on a computerized database in their 2001 study.<sup>17</sup> By using this method, the researchers could measure the absolute frequency of the motions, as well as the change in the frequency over time.<sup>18</sup> In their analysis, they found that, not only had the rate of reliability assessment risen after *Daubert*, but that the rate of excluding evidence as unreliable also was increasing.<sup>19</sup> For example, the likelihood of a reliability issue being raised was between 68% and 71% in the four years prior to *Daubert*, and 76% to 89% in the four years afterward.<sup>20</sup> Similarly, the likelihood of evidence being found unreliable went from 36%-39% to 51%-70% in the same time period.<sup>21</sup> Dixon and Gill conclude “the results...suggest that the standards for reliability tightened in the years after the *Daubert* decision. The success rate for challenges rose, encouraging an increase in the proportion of challenges that targeted reliability.”<sup>22</sup> While the effect mitigates after 1997, Dixon and Gill’s database analysis demonstrated more frequent reliability challenges after *Daubert*.

In a study from 2002, Carol Krafka and her colleagues used a survey methodology to measure how often judges reported facing a reliability motion.<sup>23</sup> While Krafka is careful to note that the survey cannot measure the absolute frequency of reliability issues arising, the survey could measure the frequency judges report reliability issues arising in their most recent expert case, and compare the results to a pre-*Daubert* survey result.<sup>24</sup> The 1998 survey finds that expert admissibility questions, when raised,

---

<sup>17</sup> DIXON & GILL, *supra* note 4, at 15-24 (methodology discussion).

<sup>18</sup> *Id.* at 25-32.

<sup>19</sup> *Id.* at Table 4.1.

<sup>20</sup> *Id.*

<sup>21</sup> *Id.*

<sup>22</sup> *Id.* at 29.

<sup>23</sup> Krafka et al., *supra* note 4, at 321.

<sup>24</sup> *Id.* at 318.

were handled more often by a motion *in limine* (72% of cases) than at trial (64%).<sup>25</sup> This contrasts greatly with survey data from 1991, when Krafska reports judges handled expert admissibility at trial in 79% of cases, while addressing it in a motion *in limine* in 32% of cases.<sup>26</sup> Krafska concludes that the data demonstrates that reliability issues are arising more frequently with a deliberative pre-trial motion than in a less-structured trial challenge.<sup>27</sup>

One final study, also from 2002, again addresses the frequency of reliability challenges. Using a database analysis technique evaluating reported cases, Jennifer Groscup and her colleagues evaluated how often *Daubert* issues arose in criminal cases.<sup>28</sup> In reviewing state and federal cases, Groscup found that over 74% of experts had been admitted at the trial court level.<sup>29</sup> At the appellate level, the rate of admittance remained very high, at 69.1%.<sup>30</sup> Groscup found that these rates did not change based on the *Daubert* shift, but that the party offering the testimony did demonstrate an effect; prosecution experts were significantly more likely to be admitted (95.8%) than defense experts (7.8%) at the trial level.<sup>31</sup>

So while Dixon and Gill, Krafska, and Groscup provide some data on the frequency of *Daubert* reliability challenges, the data is far from complete. In addition, the studies rely on surveys from 1991 and 1998<sup>32</sup> or a database of cases decided between 1980-1999 (Dixon and Gill) or 1988-1998 (Groscup).<sup>33</sup> While each of these studies is helpful, updated research could help illuminate current trends in reliability challenges.<sup>34</sup>

---

<sup>25</sup> *Id.* at 321.

<sup>26</sup> *Id.*

<sup>27</sup> *Id.* at 330.

<sup>28</sup> Groscup et al., *supra* note 4, at 342-44 (database methodology discussed).

<sup>29</sup> *Id.* at 345.

<sup>30</sup> *Id.*

<sup>31</sup> *Id.* at 346.

<sup>32</sup> Krafska et al., *supra* note 4, at 311-15.

<sup>33</sup> DIXON & GILL, *supra* note 4, at 15-24; Groscup et al., *supra* note 4, at 342-44.

<sup>34</sup> Many commentators also have noted the lack of specific data in this area. See, e.g., Margaret A. Berger, *Expert Testimony Trends in State Practice and the Uniform Rules of Procedure*, SM060 ALI-ABA 677 (2007) (mentioning lack of data on *Daubert* motions); Andrew I. Gavil, *Defining Reliable Forensic Economics in the Post-Daubert/Kumho Tire Era: Case Studies from Antitrust*, 57 WASH. & LEE L. REV. 831, 877 (2000) ("[I]t is critical to know how often *Daubert* motions are being filed and granted and in what sorts of cases. . . . A survey of district court judges, as was done in

## 2. Procedural Methodology of Handling Gatekeeping

Prior studies did not just analyze the frequency of reliability motions, but they also measured the procedures used by judges to decide them. In this area, survey research was particularly important in measuring how judges report their handling of the motions.

Krafka's 2002 study evaluated data from a 1998 survey of federal judges, which asked them about the methods used in court addressing expert issues. Krafka asked judges about many methods, including: clarifying questions from the bench, expert reports under Fed. R. Civ. P. 26, *Daubert* hearings, pretrial conferences under Fed. R. Civ. P. 16, independent experts, special masters under Fed. R. Civ. P. 53, and others.<sup>35</sup> Judges used some of the techniques, such as questions from the bench, *Daubert* hearings, or pretrial conferences, in any case involving an expert.<sup>36</sup> The judges reserved other methods for complex expert cases; these techniques included: independent experts, special masters, or requesting instruction from the parties on their expertise.<sup>37</sup> The survey responses of federal judges indicate a wide range of methodologies for handling expert issues, including both those suggested by Justice Breyer in *Joiner* and others not specifically endorsed.

State court judges handle reliability inquires as well. Shirley Dobbin and her colleagues reported on state court judges' responses to a survey in a 2007 study. Dobbin found that state court judges are less likely to use clarifying questions from the bench (69% would use).<sup>38</sup> State court judges were also less likely to ask for special instruction from the parties (29% would use, vs.

---

the Rule 11 context, might provide instructive."). To review the 2012 survey results, see *infra* Part III.B.i.

<sup>35</sup> Krafka et al., *supra* note 4, at Table 5.

<sup>36</sup> *Id.* (percentage of judges that used the procedure in any case with an expert: questions from the bench – 88.2%; *Daubert* hearing – 48.6%; pretrial conferences – 53.5%). All of these exceed the percentage that used the technique for cases with complex expert testimony. *Id.*

<sup>37</sup> *Id.* (percentage using the procedure in complex expert cases: independent experts – 16.2%; special masters – 16.6%; requesting instruction from the parties – 28.1%). All of these exceed the percentage that used the technique in any case with an expert. *Id.*

<sup>38</sup> Dobbin et al., *supra* note 4, at Table 3. This compares to 94% of federal judges; Krafka et al., *supra* note 4, at Table 5.

39% of federal judges).<sup>39</sup> However, state judges are more likely to use an independent expert, since 36% would do so (compared to 26% of federal judges).<sup>40</sup> Results on other factors are mixed.

Krafka and Dobbin provide a snapshot of the methodology of handling expert reliability motions, with both state and federal court judges. As with the research on frequency of reliability issues arising, the surveys rely on data from the 1990s.<sup>41</sup> Again, while each of these studies is helpful, updated research could help illuminate current trends in reliability challenges.<sup>42</sup>

### *B. Substantive Factors Used to Evaluate Reliability*

Beyond how often and by what methods judges evaluate reliability, empirical research has also evaluated the substantive factors judges use in deciding expert motions. *Daubert* itself suggested five factors for judges to consider, but research would be needed to decide if these factors are being used and which are most helpful.

Using statistical tools to analyze their database of published cases, Dixon and Gill could measure how often the *Daubert* factors appeared in reported case decisions.<sup>43</sup> Not surprisingly, after 1993's *Daubert* decision, the incidence of the five factors increased dramatically. Dixon and Gill found a statistically significant rise in the mention of the factors in the years after 1993, particularly in the period two-to-four years after *Daubert*.<sup>44</sup> After 1997, the incidence of the factors appearing lessens, perhaps due to a more experienced judicial gatekeeper narrowly focusing on the contested issue.<sup>45</sup> Finally, Dixon and Gill found the two factors of

---

<sup>39</sup> Dobbin et al., *supra* note 4, at Table 3; Krafka et al., *supra* note 4, at Table 5.

<sup>40</sup> *Id.*

<sup>41</sup> Dobbin et al., *supra* note 4, at 3 (state court surveys from 1999); Krafka et al., *supra* note 4, at 311 (federal court surveys from 1998 and 1991).

<sup>42</sup> To review the 2012 survey results, see *infra* Part III.

<sup>43</sup> DIXON & GILL, *supra* note 4, at xiii-xiv, 15. Note that these database samples are completely different from those used in the studies conducted by Groscup, as Dixon and Gill used solely civil cases and Groscup used solely criminal cases. *Cf.* Groscup et al., *supra* note 4, at 344.

<sup>44</sup> DIXON & GILL, *supra* note 4, at 39-40.

<sup>45</sup> *Id.* at 40-41.

general acceptance and peer review remain most relevant to judicial inquiries in the last measured period, in that order.<sup>46</sup>

Groscup's work mirrors this result about relative importance of the factors, again using analysis of published opinions from a major commercial database.<sup>47</sup> When analyzing the importance of the *Daubert* factors, Groscup found that the importance of all factors was low in comparison with other non-factor considerations such as relevance, reliability, or expert qualifications.<sup>48</sup> But when comparing solely the *Daubert* factors, the two factors that seem to be most important in assessing expert admissibility are general acceptance and peer review.<sup>49</sup> Just as with Dixon and Gill, the two factors also can be ordered as general acceptance as the most helpful, and peer review as the second most helpful.<sup>50</sup> As a result, both Dixon and Gill and Groscup have found, using statistical assessment of different opinions in a computerized database, that error rate and falsifiability are the less-helpful factors.

Survey data support these findings as well. In her 2001 study, Sophia Gatowski and her research group surveyed 400 state court judges about scientific evidence and gatekeeping.<sup>51</sup> Of the four factors, Gatowski also found general acceptance to be the most likely *Daubert* factor to be used by a judge in assessing reliability.<sup>52</sup> Of all judges surveyed, 93% found general acceptance useful in assessing the merits of scientific evidence, with 64% finding the factor very useful.<sup>53</sup> Slightly fewer judges, at 92%, felt that peer review was useful at all, with 52% finding it was very useful.<sup>54</sup> Error rate at 91% (and 54% very useful) and falsifiability at 88% (38% very useful) were less often cited as critical to assessing expert evidence.<sup>55</sup>

---

<sup>46</sup> *Id.* at 39.

<sup>47</sup> *See supra* note 43 (note that the two databases are completely different).

<sup>48</sup> Groscup et al., *supra* note 4, at Table 5.

<sup>49</sup> *Id.*

<sup>50</sup> *Id.*

<sup>51</sup> Gatowski et al., *supra* note 4, at 441.

<sup>52</sup> *Id.* at 446.

<sup>53</sup> *Id.* at 447.

<sup>54</sup> *Id.*

<sup>55</sup> *Id.* at 444-45.

Just as with the studies on the frequency of reliability assessment and procedural methods of handling reliability issues, the data on substantive factors is very helpful but somewhat dated. Updated research in the area would be important to assess current trends in judicial management of these reliability motions.

## II. RELIABILITY MOTIONS AND SUMMARY JUDGMENT: COMMENTARY & PRIOR STUDIES

Considering the increased scrutiny placed on expert evidence due to the gatekeeping role and the increased likelihood of challenges to experts after *Daubert*, many commentators have suggested a relationship between reliability screening and summary judgment.<sup>56</sup>

In the year after the *Daubert* decision, Joseph Sanders suggested that judges who perform gatekeeping on expert evidence often blur the line between the question of admissibility and the issue of sufficiency of the evidence.<sup>57</sup> Sanders suggested that judges have become, in cases involving complex scientific evidence, more likely to restrict admissibility, based on both judicial efficiency and jury competence grounds.<sup>58</sup> As a result, he believed that judges were using admissibility challenges for case management ends.<sup>59</sup> Since admissibility challenges are more likely after *Daubert*, Sanders's work suggests that reliability analysis will inevitably lead to sufficiency screening, thereby increasing summary judgment.<sup>60</sup>

Sanders is not alone in discussing this concern. Neil Vidmar, also writing in the year after *Daubert*, also highlighted the

---

<sup>56</sup> This is not to say judges would not avoid juries prior to *Daubert*, since studies suggested as much earlier than 1993. Joe S. Cecil, Valerie P. Hans & Elizabeth C. Wiggins, *Citizen Comprehension of Difficult Issues: Lessons from Civil Jury Trials*, 40 AM. U. L. REV. 727, 737-40 (1991) (skepticism of juries leads to judicial involvement in factual issues). But the concern of judicial involvement in the facts, underlying the Cecil study for example, is clearly raised by the mechanism of gatekeeping post-*Daubert*.

<sup>57</sup> Sanders, *supra* note 7, at 1434.

<sup>58</sup> *Id.* at 1429.

<sup>59</sup> *Id.* at 1434.

<sup>60</sup> Sanders states this point again in a later article, Joseph Sanders, *The Merits of the Paternalistic Justification for Restrictions on the Admissibility of Expert Evidence*, 33 SETON HALL L. REV. 881, 930, 939 (2003) (discussing DIXON & GILL, *supra* note 4; finding that judges use reliability partially to avoid the jury).

concern that judges may be using reliability screening to assess the merits of the underlying dispute.<sup>61</sup> Vidmar suggested that one of the justifications for that sufficiency screening – the perception of juror incompetence with complex testimony<sup>62</sup> – may not be so clear as to justify judicial intervention in these cases.<sup>63</sup> Yet the underlying concern remains similar to Sanders; judges may overstep the bounds of the issue of reliability to judge the sufficiency of the evidence in a manner similar to summary judgment.

Multiple other commentators have discussed this concern. In a 1998 article, Peter Schuck directly tied reliability decisions to increases in summary judgment. Schuck theorizes that evidence exclusion in complex cases, pursuant to *Daubert* gatekeeping, “have produced some blanket exclusions of evidentiary categories, which can then justify courts in using directed verdicts or motions for summary judgment in favor of defendants to avoid jury trials.”<sup>64</sup> Lucinda Finley, writing in 1999, agrees with this concern:

Some judges have used their gatekeeper power to wade into disputed scientific territory to try to resolve or choose sides in scientific controversies. Ignoring the usual legal rule that decisions about the admissibility of evidence are meant to be distinct from decisions about the weight or sufficiency of [the] evidence, these courts have adopted the legal position that *there is no difference under Daubert-Joiner between the sufficiency and admissibility determinations.*<sup>65</sup>

Finally and more recently, Arthur Hellman has described this same mechanism when discussing litigants’ concern with

---

<sup>61</sup> Neil Vidmar, *Are Juries Competent to Decide Liability in Tort Cases Involving Scientific/Medical Issues? Some Data from Medical Malpractice*, 43 EMORY L.J. 885, 888-89 (1994).

<sup>62</sup> Sanders has offered this as one of the two justifications used for the judges’ involvement in sufficiency screening, along with judicial efficiency. Sanders, *supra* note 7, at 1429. See also *Developments in the Law: The Civil Jury*, 110 HARV. L. REV. 1408, 1489-93 (1997).

<sup>63</sup> Vidmar, *supra* note 61, at 910-11.

<sup>64</sup> Schuck, *supra* note 7, at 498.

<sup>65</sup> Finley, *supra* note 7, at 347-48 (emphasis added).

appearing in federal court.<sup>66</sup> Hellman quotes a prominent litigator who blames *Daubert's* codification of a skeptical attitude toward juries, along with the increased incidence of summary judgment, resulting in plaintiffs fleeing federal court and defendants rushing to get federal jurisdiction.<sup>67</sup> Hellman concludes that the perception that federal judges will more aggressively use summary judgment is “widely shared” among litigants.<sup>68</sup>

Commentators like Sanders, Vidmar, Schuck, Finley, and Hellman suggest that *Daubert* is a way for judges to assess not just the reliability, but also sufficiency, of evidence. But this proposition, repeated as though it may be, should be empirically studied to test whether the connection can be empirically verified.<sup>69</sup>

Before discussing my study evaluating this issue, it is important to note that many of the same empirical research gaps exist in the summary judgment area as exist in expert reliability area. For example, in trying to measure the frequency of summary judgment, research is quite varied. Some studies show summary judgment rates increasing over time. Dixon and Gill's study, reviewing many aspects of expert reliability as discussed *supra*, also reviewed summary judgment before and after *Daubert*.<sup>70</sup> In their findings, the prevalence of summary judgment rose after *Daubert*, as did the frequency of judges granting the motion.<sup>71</sup> On the other hand, Dixon and Gill suggest this may or may not be directly tied to *Daubert*, and suggest further study.<sup>72</sup>

Similarly, in a 2007 study Joe Cecil and his research group addressed the prevalence of summary judgment in federal courts over time.<sup>73</sup> Cecil used a sample of approximately 15,000 federal cases between 1975 and 2000, in six different federal districts, to find the rates of motions and granting of summary judgment.<sup>74</sup>

---

<sup>66</sup> Arthur Hellman, *Another Voice For the “Dialogue”: Federal Courts as a Litigation Course*, 53 ST. LOUIS U. L.J. 761, 766-67 (2009).

<sup>67</sup> *Id.* at 767.

<sup>68</sup> *Id.*

<sup>69</sup> Regarding the 2012 survey results, see *infra* Part III.

<sup>70</sup> DIXON & GILL, *supra* note 4, at 55-57.

<sup>71</sup> *Id.* at Table 7.3.

<sup>72</sup> *Id.* at 57.

<sup>73</sup> Cecil et al., *supra* note 8, at 861-62.

<sup>74</sup> *Id.* at 877.

The results demonstrate a small increase in the frequency of motions for summary judgment, from 12% in 1975 to 21% in 2000.<sup>75</sup> The frequency of a judge granting summary judgment rose from 6% of cases in 1975 to 12% in 2000, as well as the likelihood that the summary judgment terminated the case.<sup>76</sup> Cecil is quick to caution, however, that these changes are varied between courts, location, and types of cases.<sup>77</sup> As a result, the effect of *Daubert* on summary judgment cannot be assessed by this study alone.

On the other hand, several other studies show no broad-based changes with summary judgment. In 2007, Theodore Eisenberg and Charlotte Lanvers evaluated rates of summary judgment using reported cases from three periods, 1975-76, 1980-81, and 2001-02.<sup>78</sup> When they did so, they found no clear changes in summary judgment between those periods.<sup>79</sup> Having reviewed these data, Eisenberg and Lanvers concluded that “we find no evidence of a broad-based increase in summary judgment after [the 1986 trilogy expanding summary judgment].”<sup>80</sup> Instead, they conclude the variation in rates is usually just between districts, or by case types, more than any other factor.<sup>81</sup>

Other studies do not measure changes over time, but instead look at factors affecting summary judgment within a single year. In their Federal Judicial Center analysis, Joe Cecil and George Cort measured the frequency of summary judgment from all of the different federal circuits in 2006.<sup>82</sup> Their work relied on reporting

---

<sup>75</sup> *Id.* at 882, 896. Even so, the frequency with which litigants file for summary judgment has remained static since 1986. *Id.* at 882.

<sup>76</sup> *Id.* at 882-83.

<sup>77</sup> *Id.* at 883. Indeed logistic regression appears to show varied results based on type of case, whether the motion was filed, granted, or resulted in termination of the case, and location. *Id.* at 883-86, 890-96.

<sup>78</sup> Eisenberg & Lanvers, *supra* note 8, at 9-11.

<sup>79</sup> *Id.* at 19.

<sup>80</sup> *Id.* The summary judgment trilogy includes three cases decided in 1986: *Matsushita Elec. Indus. Co. v. Zenith Radio Corp.*, 475 U.S. 574 (1986), *Anderson v. Liberty Lobby, Inc.*, 477 U.S. 242 (1986), and *Celotex Corp. v. Catrett*, 477 U.S. 317 (1986). These cases are generally perceived as re-invigorating summary judgment use in federal court. *See, e.g.*, Cecil et al., *supra* note 8, at 861-62; Eisenberg & Lanvers, *supra* note 8, at 1. For a summary of commentary on this issue, see Paul Mollica, *Federal Summary Judgment at High Tide*, 84 MARQ. L. REV. 141, 153 n.69 (2000).

<sup>81</sup> Eisenberg & Lanvers, *supra* note 8, at 19.

<sup>82</sup> Memorandum from Joe Cecil & George Cort, Fed. Judicial Ctr., to Hon. Michael Baylson for Estimates of Summary Judgment Activity in Fiscal Year 2006 (June 15,

of cases through the federal court database, including a total of nearly 180,000 cases.<sup>83</sup> When they reviewed the database, they found summary judgment motions occurring in 17 out of 100 cases that were terminated, and granted – either in whole or in part – in 60% of cases.<sup>84</sup> Cecil and Cort did specifically caution that summary judgment rates vary significantly between different circuits or by case type.<sup>85</sup>

Just as with the analyses of reliability issues, studies of summary judgment result in mixed conclusions. As with reliability then, summary judgment is an area ripe for additional updated empirical analysis.

### III. A NEW SURVEY ON JUDICIAL HANDLING OF GATEKEEPING

Considering the coverage and age of prior research on expert reliability motions, I decided to empirically study the frequency of, procedural methodology for, and substantive factors relevant to deciding challenges to expert reliability. Considering the connection made between reliability and summary judgment, and inconsistent data regarding summary judgment motions, I also decided to analyze summary judgment. This study also assesses both the frequency and handling of summary judgment motions, and compares those motions to reliability motions.

#### *A. Study Methodology*

To these ends, I created a survey for state court trial judges, who would be very familiar with both motions.<sup>86</sup> Not only could the survey help establish the methodologies of the two motions, but in so doing, I could also analyze similarities between the motions in the responses of the judges.

Selection of states to use for the study involved several important factors. Initially, I wanted to ensure a diverse group of states were involved in order to get the broadest-based sample

---

2007), available at [http://www.fjc.gov/public/pdf.nsf/lookup/sujufy06.pdf/\\$file/sujufy06.pdf](http://www.fjc.gov/public/pdf.nsf/lookup/sujufy06.pdf/$file/sujufy06.pdf).

<sup>83</sup> *Id.* at 1.

<sup>84</sup> *Id.* at 2-3.

<sup>85</sup> *Id.* at 1-2.

<sup>86</sup> This study was approved by the Drake University Institutional Review Board (IRB) on May 22, 2012.

possible. I also wanted to pick states that had different standards for determining the reliability of scientific evidence, so that some judges would have *Daubert* as their home standard and others would not.<sup>87</sup> This would allow for analysis of variations in handling motions based on the home-state standard.<sup>88</sup> States were then grouped by region, and for each region I needed a state that had a *Daubert* admissibility standard and one that retained the *Frye* standard so that I could cross-compare different factors.<sup>89</sup> Finally, each state would need to have similar evidentiary and procedural rules in the relevant areas including: independent experts, pretrial conferences, questions from the bench, and special masters. With all these considerations, choosing states involved substantial background research.

Based on an initial analysis of potential states, I chose the following states for the survey: Alabama, Mississippi, New Mexico, Arizona, Minnesota, and Michigan. The sample included three states that have adopted the *Daubert* standard – Mississippi,<sup>90</sup> New Mexico,<sup>91</sup> and Michigan<sup>92</sup> – and three states that have continued using the *Frye* standard – Alabama,<sup>93</sup> Arizona,<sup>94</sup> and

---

<sup>87</sup> The alternative standard used by many states is the *Frye* “general acceptance” standard, which had been cast aside in federal court in *Daubert*. *Daubert v. Merrell Dow Pharm.*, 509 U.S. 579, 589 (1993) (FED. R. EVID. 702 provides the standard for expert admissibility, rather than “general acceptance” standard); *Frye v. United States*, 293 F. 1013 (D.C. Cir. 1923) (general acceptance standard adopted by a federal circuit court). *See also supra* note 15 and accompanying text.

<sup>88</sup> Regarding the 2012 survey results in this area, see Part III.B.iv. *See also* Andrew W. Jurs, *Questions from the Bench and Independent Experts: A Study of the Practices of State Court Judges*, 74 U. PITT. L. REV. 47, 65 (2012) (analysis of judges’ state admissibility standard and their use of evidentiary techniques in court).

<sup>89</sup> The analysis of these groupings is beyond the scope of this paper, and will be presented in a separate work. Andrew W. Jurs, *Gatekeeper with a Gavel II*, (unpublished manuscript)(on file with author). *See also infra* note 108.

<sup>90</sup> *Miss. Transp. Comm’n v. McLemore*, 863 So. 2d 31, 35 (Miss. 2003) (*Daubert* standard adopted in State of Mississippi, to replace previously-used *Frye* standard).

<sup>91</sup> *State v. Alberico*, 861 P.2d 192, 203 (N.M. 1993) (*Daubert* standard adopted in state court).

<sup>92</sup> *Gilbert v. DaimlerChrysler Corp.*, 685 N.W.2d 391, 408 (Mich. 2004) (recognizing change to *Daubert* standard by a 2003 amendment to MICH. R. EVID. 702).

<sup>93</sup> *Bagley v. Mazda Motor Corp.*, 864 So. 2d 301, 310 (Ala. 2003) (*Frye* standard used in Alabama; only exception is DNA cases by ALA. CODE §36-18-30); *Courtaulds Fibers, Inc. v. Long*, 779 So. 2d 198, 202 (Ala. 2000) (same).

<sup>94</sup> *Logerquist v. McVey*, 1 P.3d 113, 134 (Ariz. 2000) (“[W]e retain the *Frye* rule but continue to apply it as described in *Hummert*. We reject the *Joiner/Kumho* interpretation of FED. R. EVID. 702 and continue to apply ARIZ. R. EVID. 702 as written

Minnesota.<sup>95</sup> The states also break into three distinct regions – South (AL & MS), West (AZ & NM), and Midwest (MN & MI) – each with one *Daubert* and one *Frye* state. Finally, each state has similar rules of evidence and procedure in the important areas of summary judgment,<sup>96</sup> independent experts,<sup>97</sup> questions from the bench,<sup>98</sup> pretrial conferences,<sup>99</sup> and special masters.<sup>100</sup> The states selected offered the best available combination of the important factors.<sup>101</sup>

---

and interpreted by our cases.”). The *Frye* standard in use in Arizona changed to *Daubert* by an amendment to Rule 702, effective January 1, 2012. ARIZ. R. EVID. 702 (2012). Considering that judges were asked about the entirety of their bench experience, and the survey was completed by judges only months after the January 1, 2012 change, I believe categorization of Arizona as a *Frye* state remained appropriate.

<sup>95</sup> *Goeb v. Tharaldson*, 615 N.W.2d 800, 814 (Minn. 2000) (*Frye* standard remains in Minnesota).

<sup>96</sup> ALA. R. CIV. P. 56 (2012); ARIZ. R. CIV. P. 56 (2012); MICH. CT. R. 2.116; MINN. R. CIV. P. 56.01-56.07 (2006); MISS. R. CIV. P. 56; N.M. DIST. CT. R. CIV. P. 1-056.

<sup>97</sup> ALA. R. EVID. 706; ARIZ. R. EVID. 706; MICH. R. EVID. 706; MINN. R. EVID. 706; MISS. R. EVID. 706; N.M. R. EVID. 11-706.

<sup>98</sup> ALA. R. EVID. 614; ARIZ. R. EVID. 614; MICH. R. EVID. 706; MINN. R. EVID. 614; MISS. R. EVID. 614; N.M. R. EVID. 11-614.

<sup>99</sup> ALA. R. CIV. P. 16 (2011); ARIZ. R. CIV. P. 16; MICH. CT. R. 2.401; MINN. R. CIV. P. 16.01; MISS. R. CIV. P. 16; N.M. DIST. CT. R. CIV. P. 1-016.

<sup>100</sup> ALA. R. CIV. P. 53 (2011); ARIZ. R. CIV. P. 53; MINN. R. CIV. P. 53.01-53.09; MISS. R. CIV. P. 53; N.M. DIST. CT. R. CIV. P. 1-053. The single greatest difference between states in the study involves the rule on special masters in Michigan. MICH. CT. R. 3.913. Under this rule, a special master may assist in juvenile proceedings, but there is not a general rule for special masters. The extent to which appointment of a special master can be appointed under implicit authority was debated in detail in *Oakland County Prosecutor v. Beckwith*, in which the court simultaneously expressed belief that implicit powers should exist, while denying the special master due to specific case precedent. 619 N.W.2d 172, 174-75, 177 (Mich. 2000) (citing *Carson Fisher Potts & Hyman v. Hyman*, 559 N.W.2d 54 (Mich. Ct. App. 1996) (special masters impermissible delegation of judicial authority)). Considering the points of agreement in all other areas, I decided to proceed with the use of Michigan in the survey. It remained much closer in rules to Minnesota than any alternative state pairing available. *See infra* note 101 (explaining problems with other state pairings).

<sup>101</sup> Considering the many different areas of consideration, I did attempt to determine if any alternative state pairing would be a better option than those states chosen. One possible pairing option was Oregon (*Daubert*) and Washington (*Frye*). The problem here was that while Washington’s Rules of Evidence and Civil Procedure are similar to the other states, Oregon’s are not. Oregon lacks specific rules on independent experts and questions from the bench, and while there is authority for an inherent power, its extent is unclear. *See State v. Mains*, 669 P.2d 1112, 1123-24 (Or. 1983) (inherent authority exists, but questioning is to be rare). In addition, the state civil rule regarding pretrial conferences appears to be significantly different from the federal rule and all other states in the survey. ORE. UNIFORM TRIAL COURT RULE 6.010;

For each of the states selected, I anticipated sending surveys to trial court judges. However, I limited the judges eligible to participate to only those judges who serve on the highest-level trial court within the state.<sup>102</sup> State judiciary websites provided lists of currently-serving judges, and provided the final lists of eligible judges.<sup>103</sup> The initial survey group included a total of 996 participants.

Every judge meeting the sample qualifications received an initial letter explaining the survey and its general goals, along with an explanation of how to participate. Judges were invited to

---

*compare with* FED. R. CIV. P. 16 and *supra* note 99. The limited scope of the state rule appears to be one reason for local rules on the same issue. *See, e.g.*, CLATSOP COUNTY LOCAL RULE 6.011 (2012); MULTNOMAH COUNTY SUPPLEMENTARY LOCAL RULE 6.014 (2012), *available at* <http://courts.oregon.gov/Multnomah/docs/CourtRules/SLR2012EffectiveFebruary1.2012.pdf>. One other possible state pairing involved Maryland (*Frye*) and West Virginia (*Daubert*). While their rules are substantially similar – West Virginia’s rule on special masters being one difference – the size and makeup of the judiciary of each state merited second thoughts. In Maryland, there are over 250 judges who would have been included, while in West Virginia there are 65. Added to the potential dissimilarity between the states’ urban-rural makeup, which is another judicial category for potential analysis, the differences seemed too stark. *See infra* note 108. No other *Daubert* state exists to pair with Maryland, since Pennsylvania follows *Frye*; Virginia has a state-specific admissibility standard, and New Jersey has rejected wholesale adoption of *Daubert* and retains *Frye* standards. Alice B. Lustre, Annotation, *Post-Daubert Standards for Admissibility of Scientific and Other Expert Evidence in State Courts*, 90 A.L.R.5th 453, pp. 73, 76, 544 (2001) (citing *State v. Chun*, 943 A.2d 114 (N.J. 2008); *Grady v. Frito-Lay*, 839 A.2d 1038 (Pa. 2003); *Spencer v. Com.*, 393 S.E.2d 609 (Va. 1990)). As for North Carolina, the state appeared to adopt *Daubert* in 1995, but in a later case backtracked away from wholesale adoption. *Howerton v. Arai Helmet, Ltd.*, 597 S.E.2d 674, 689 (N.C. 2004) (backtracking); *State v. Goode*, 461 S.E.2d 631 (N.C. 1995). With the lack of clarity on the issue since *Howerton* in 2004, the state did not seem to be an appropriate choice either.

<sup>102</sup> This ensures both that they are using the state rules of evidence but also that they have the authority to hear the highest-level civil claims, which could involve complex evidence.

<sup>103</sup> Each list of state court judges was obtained from the following websites: List of Courts and Administrative Officers, ALABAMA STATE BAR, [http://www.alabar.org/olDirectory/index\\_judicial.cfm](http://www.alabar.org/olDirectory/index_judicial.cfm) (last visited Oct. 14, 2013); Arizona Courts Locator, ARIZONA SUPREME COURT, <http://www.azcourts.gov/AZCourts/AZCourtsLocator.aspx> (last visited Oct. 14, 2013); Michigan Courts Directories, MICHIGAN COURTS, <http://courts.mi.gov/Self-Help/Directories/Pages/Trial-Court-Directory.aspx> (last visited Oct. 14, 2013); Judicial Directory, MINNESOTA JUDICIAL BRANCH, <http://www.mncourts.gov/?page=25> (last visited Oct. 14, 2013); Judiciary Directory, STATE OF MISSISSIPPI JUDICIARY, <http://courts.ms.gov/judiciarydirectory/judiciarydirectory.html> (last visited Oct. 14, 2013); Court Sites, NEW MEXICO ADMINISTRATIVE OFFICE OF THE COURTS, <http://www.nmcourts.gov/othercourts.php> (last visited Oct. 14, 2013).

log in to a secure website using a personal “authenticator” supplied by the researcher.<sup>104</sup> After some background questions, the online survey asked judges about motions challenging expert reliability, including their frequency, procedures used to decide them, substantive factors used to determine them, and overall comfort level with the motions.<sup>105</sup> The survey asked additional questions of judges who have limited expert testimony due to a motion challenging expert reliability, regarding the extent of the limit, the reasons for the limit, and the judge’s comfort level in limiting the expert testimony. In addition to the questions on reliability motions, the survey asked a series of similar questions about motions for summary judgment. Each judge was asked about the frequency of the motion and the procedures used to decide those motions.<sup>106</sup> As with reliability motions, the survey asked additional questions of judges who have granted – in full or in part – motions for summary judgment, to understand the frequency of a judge granting the motion and the judge’s comfort level in doing so.

Following an initial round of letters, I contacted the judges who had yet to complete the survey to remind them of the survey and ask for their participation. Following the second round of participation, the Study includes a final survey group of 158 judges.<sup>107</sup> Upon finalizing the response data, I then analyzed the data from all judges to determine how judges handle reliability and summary judgment motions.

### *B. Analysis of Responses*

The data from all judges provides the first insights from the survey on judicial management of reliability motions and their relationship to summary judgment. To review the data, I will first

---

<sup>104</sup> The survey used the Qualtrics online survey system, which allows online responses and collects the data. The authenticator was created using a random word generator and consisted of a four to six-letter word unique to each judge.

<sup>105</sup> The complete survey instrument is on file with the author.

<sup>106</sup> While perfect symmetry between the handling of reliability motions and the handling of summary judgment would be ideal, the substantive consideration in granting summary judgment is quite clear by rule. *See, e.g.*, FED. R. CIV. P. 56. No questions regarding the substantive considerations in granting summary judgment were needed.

<sup>107</sup> My thanks to all judges who took the time to participate.

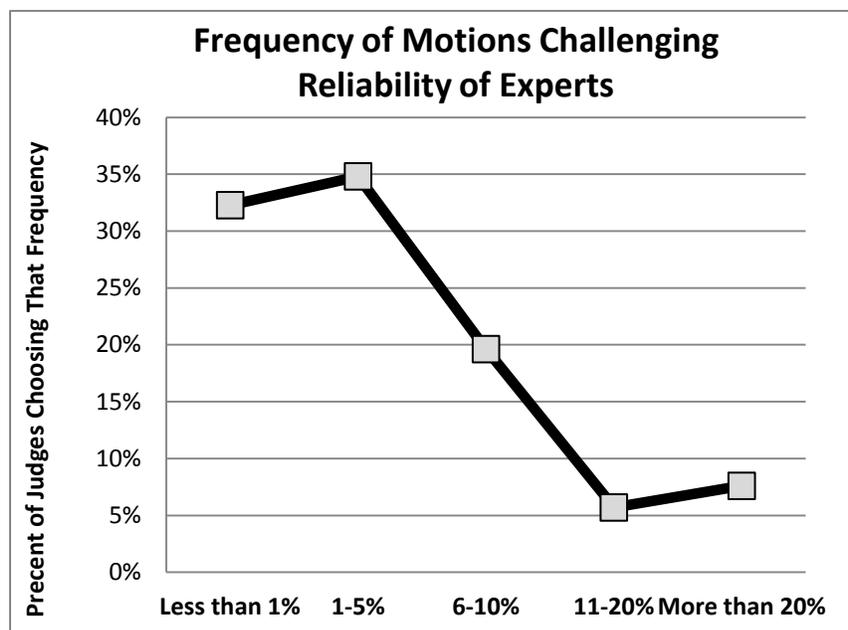
evaluate the frequency of reliability motions, then the procedural methods of handling the motions and the substantive factors used to decide them. I will then finish the analysis of reliability motions by reviewing how often, why, and to what extent judges grant motions challenging expert reliability. Then I will switch to discuss survey data on the judicial handling of motions for summary judgment, reviewing their frequency, procedural methods of reviewing them, and how often they are granted. I will finish the analysis of all judicial responses by comparing the data about reliability motions to summary judgment, and whether the closeness of the connection can be empirically verified. In addition to the general analysis, I also will analyze variations in responses based on the judge's home-state admissibility standard, to see if there are statistically significant differences in the way judges handle motions due to differences in substantive law.<sup>108</sup>

### *C. Judicial Handling of Reliability Motions*

Judges taking the survey answered a variety of questions discussing reliability motions challenging expert testimony, starting with questions about their frequency. The first question on frequency of motions asked judges: "In what percentage of cases with expert witnesses do you see a motion challenging the reliability of expert testimony?" A substantial majority of judges believed such motions occurred in less than 5% of cases with experts, with 32% answering that the motion occurred in less than 1% of cases ( $n=51$ ), and an additional 35% answering that the motion occurred between 1% and 5% of the time ( $n=55$ ). This compares to a total of 33% who believed the motions were in 6% or more of cases with experts ( $n=52$ ).

---

<sup>108</sup> The analysis of judges with different home-state standards is presented *infra* Part III.F. In addition to the home-state standard analysis, the data will also allow analysis of responses by many other demographic groupings, including: state, region, location of the court (urban/rural), experience with and comfort with math and science, years of experience as a judge, years in practice, and practice side/subject. The analysis of these groupings is beyond the scope of this paper, and will be presented in a separate work. Jurs, *supra* note 89 (work in progress).



**Figure 1: Frequency of Motions Challenging Expert Reliability, in All Cases with Experts**

After asking about the frequency with which judges see reliability motions, the survey asked all judges what substantive factors they believe are helpful in deciding the motion. The possible responses included the “*Daubert* factors” of: technique can and has been tested, subjected to peer review and publication, known or potential rate of error, existence of standards controlling the technique, and general acceptance.<sup>109</sup> Judges were instructed to choose any factors they believed were helpful, and were also given the option of choosing an open-ended “other” response and then being permitted to enter text to explain their answer.

Judicial responses are recorded below in Figure 2, with the most helpful factors being general acceptance and testing while the least helpful factor was rate of error.

<sup>109</sup> These are the factors endorsed by the Supreme Court in *Daubert* for addressing reliability motions. See *supra* note 14 and accompanying text.

	Number Who Believe the Factor is Helpful	Percentage
Technique Can And Has Been Tested	143	90.5
Subjected To Peer Review And Publication	139	87.8
Known Or Potential Rate Of Error	111	70.2
Existence Of Standards Controlling The Technique	128	81.0
General Acceptance	152	96.2
Other	8	5.1
	Total = 158	

**Figure 2: Substantive Factors Judges Believe Are Helpful to Decide a Motion Challenging Expert Reliability**

In addition to the question as to which substantive factors were helpful in deciding reliability motions, the survey also asked judges about the two standards for scientific admissibility: *Frye* and *Daubert*. When asked directly which standard is the stricter one for gatekeeping, 61% of judges chose *Daubert* as the more restrictive standard ( $n=96$ ) while 39% answered *Frye* ( $n=62$ ). Considering the majority of judges in the response group serve in a *Frye* jurisdiction,<sup>110</sup> the endorsement of *Daubert* as the stricter standard is both interesting and counterintuitive.<sup>111</sup>

Following the questions on substantive factors in deciding reliability motions, the next questions asked judges if they had personally ruled on that type of motion and, if so, in how many cases. Of 158 respondents, 75% have ruled on a motion challenging reliability ( $n=118$ ). A slim majority of 51% has done so in two to five cases ( $n=60$ ), while 21% have done so in six to ten cases ( $n=25$ ). Judges chose the other options at lesser rates: one

<sup>110</sup> Regarding the breakdown of home-state admissibility standard of judges completing the survey and their perceptions of the stricter standard based on that classification, see *infra* Part III.F.

<sup>111</sup> Regarding the debate over which standard is stricter, see Andrew W. Jurs & Scott DeVito, *The Stricter Standard: An Empirical Assessment of Daubert's Effect on Civil Defendants*, 62 CATH. U. L. REV. 675 (2013) (forthcoming). In that study and based on statistical analysis of a dataset of millions of actual cases, the researchers conclude *Daubert* is the stricter standard. *Id.*

case – 11% ( $n=13$ ), eleven to twenty cases – 9% ( $n=11$ ), and over twenty cases – 7.6% ( $n=9$ ).

For those judges who have ruled on a motion challenging expert reliability, the survey then asked what procedural techniques were helpful in ruling on the motion. The possible responses to this question included the suggested methods of Justice Breyer from *Joiner*: questioning a witness from the bench, independent expert, special master, or hearings (with or without testimony presented).<sup>112</sup> Judges were again instructed to choose any factors they believed were helpful, and were also given the option of choosing an open-ended “other” response and then being permitted to enter text to explain their answer.

Judicial responses to this question are recorded below in Figure 3. Judges who have ruled on reliability motions are most likely to use a hearing with testimony to decide the motion, and least likely to use a special master under Rule 53 or an Independent Expert under Rule 706.

	Number Who Have Used the Technique	Percentage
Hearing With Testimony Presented	101	85.6
Hearing Without Testimony Presented	63	53.4
Questioning a Witness From the Bench	64	54.2
Independent Expert	9	7.6
Special Master	4	3.4
Other	10	8.4
	Total = 118	

**Figure 3: Procedural Methods Used by Judges to Decide Reliability Motions**

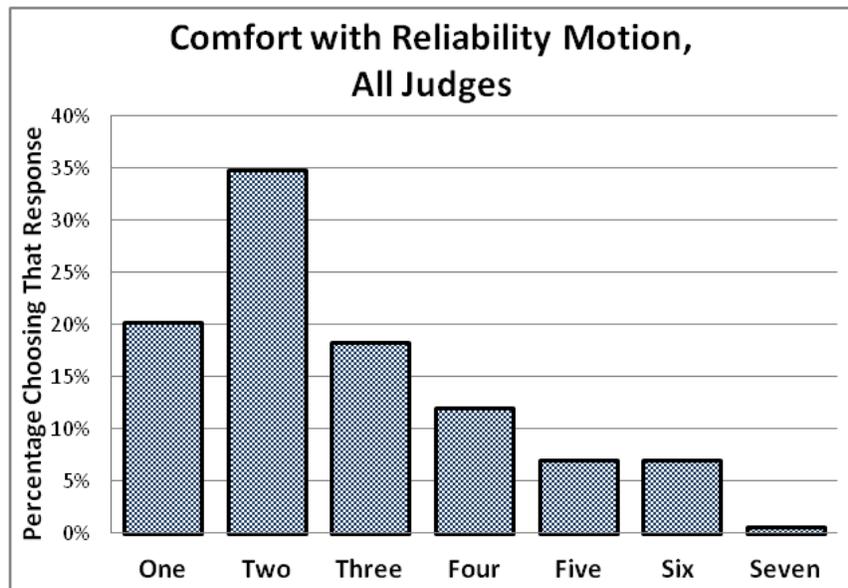
Besides asking specifically about the procedural methods and substantive considerations in deciding expert reliability motions, the survey also asked judges to rate their overall comfort with that type of motion. For this question, and for all questions asked about judicial comfort with evaluating or granting motions, the survey asked judges to rate their comfort level on a scale from one

---

<sup>112</sup> See *supra* note 15 and accompanying text.

to seven, with one being “Entirely Comfortable” and seven being “Not Comfortable.” This question was asked of all judges taking the survey.

When asked to rate their comfort level with reliability challenges on this scale, a majority (55%;  $n=87$ ) answered the question with one or two, with 20.3% at “1” and 34.8% at “2” (1:  $n=32$ ; 2:  $n=55$ ). Other answers were less frequent, with 18.4% answering “3” ( $n=29$ ), 12% answering “4” ( $n=19$ ), and 7% answering both “5” and “6” (5:  $n=11$ ; 6:  $n=11$ ). One judge answered “7” on this question.<sup>113</sup>

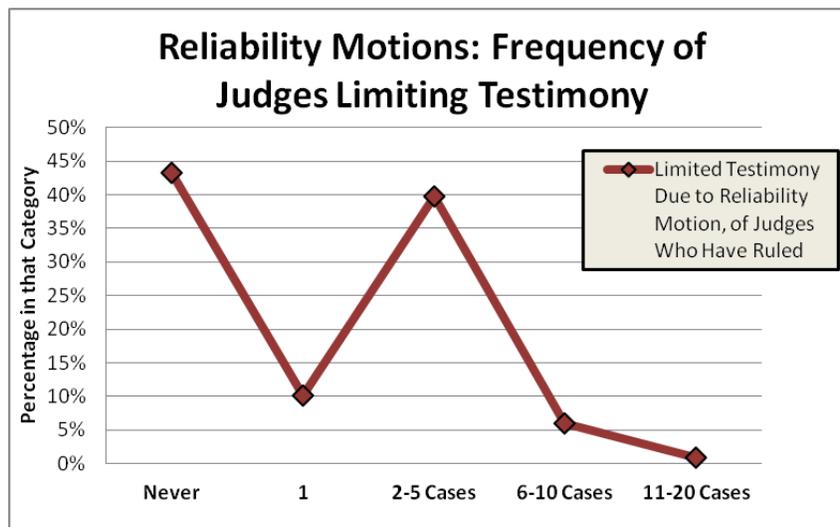


**Figure 4: Comfort Level with a Motion Challenging Reliability of Expert Evidence, All Judges**

Questions in the survey began with a series of questions about the handling of, comfort with, and considerations in deciding a motion challenging expert reliability. The next questions asked judges about whether they have granted such a motion, and if so, how often and why.

<sup>113</sup> Mean = 2.74; standard deviation = 1.49.

Of the 118 judges in the survey who have ever ruled on a reliability motion, a majority, 57%, had limited expert testimony pursuant to the motion ( $n=67$ ). Of those sixty-seven judges, most had done so fewer than five times. Twelve judges had never granted a reliability motion (18%), and 47 had granted the motion between two and five times (70%). Only eight judges had limited testimony of an expert over five times, with seven doing so between six and ten times. No judge answering the survey had limited expert testimony in over twenty cases. The variations in these responses are depicted below in Figure 5.



**Figure 5: Frequency of Limiting Testimony Pursuant to a Reliability Motion, of Judges who Have Ruled on Such a Motion**

The final group of questions on reliability motions asked those judges who had limited expert testimony about the extent of the limitation, the reasons for the limitation, and how comfortable they were with the limitation. When asked the extent of the limitation, the judges could choose among the following options: excluded all of the expert's opinions, excluded some of the expert's opinions, excluded one of the expert's opinions, otherwise limited

the opinions without excluding them, or “other” which provided the judge the chance to enter text. Judges were instructed to choose any response that applied to their experiences.

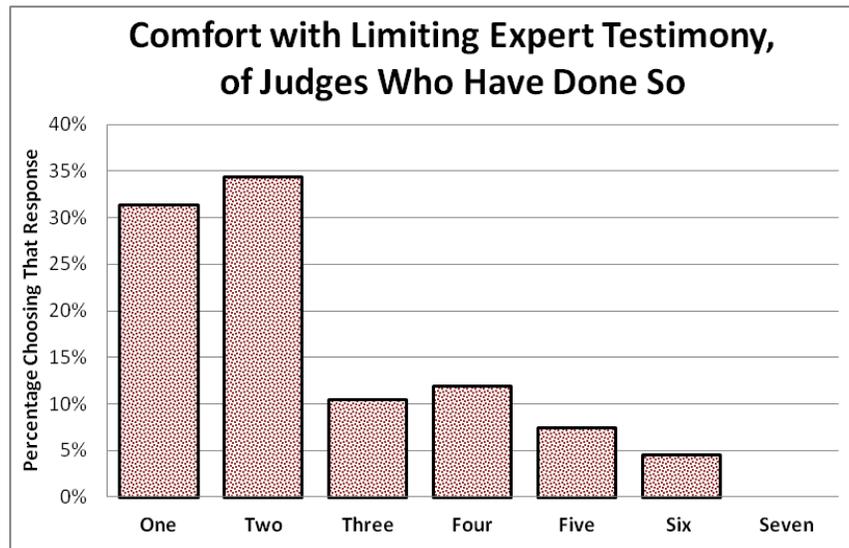
Judicial responses to this question are recorded below in Figure 6. The responses indicate that judges are limiting testimony in any number of ways, from complete exclusion to limitation without exclusion.

	Selected that Response	Percentage
Excluded All of the Expert’s Opinions	31	46.2
Excluded Some of the Expert’s Opinions	48	71.6
Excluded 1 of the Expert’s Opinions	11	16.4
Otherwise Limited the Opinions of the Expert without Excluding Them	26	38.8
Other	0	0
	Total = 67	

**Figure 6: Extent of Limitation of Expert Testimony, of Judges Who Have Limited Expert Testimony Pursuant to a Reliability Motion**

For the sixty-seven judges who had limited testimony, the survey then asked why the judge limited the testimony. Judges could select Lack of Reliability of the Underlying Scientific Principles or Lack of Reliability of the Method as Applied to the Case, and also could select “other” and enter text. As with the last question, judges could select more than one response. In response to this question, 54% of judges limited testimony due to the underlying principles being unreliable, and an identical 54% also limited testimony due to methodological unreliability ( $n=36$ , for both). Thirty-one percent of judges who have limited testimony chose “other” and entered text in response to this question ( $n=21$ ). Because a substantial portion of judges who have limited expert testimony chose “other” as the reason, additional explanation is useful to help clarify the reasons judges do so. Of the twenty-one “other” responses, a large number (15) of them cited lack of expert qualification as the basis for limiting the expert testimony.

When asked to rate their comfort level with limiting expert testimony on the one-to-seven scale, a majority (66%;  $n=44$ ) answered the question with one or two, with 31.3% at “1” and 34.3% at “2” (1:  $n=21$ ; 2:  $n=23$ ). Other answers occurred less frequently, with 10.4% answering “3” ( $n=7$ ), 11.9% answering “4” ( $n=8$ ), 7.4% answering “5” ( $n=5$ ), and 4.5% answering “6” ( $n=3$ ). No judge in the survey responded with a “7” to this question.<sup>114</sup>



**Figure 7: Comfort Level with Limiting Expert Testimony Pursuant to a Motion Challenging Reliability, Of Judges Who Have Limited Testimony**

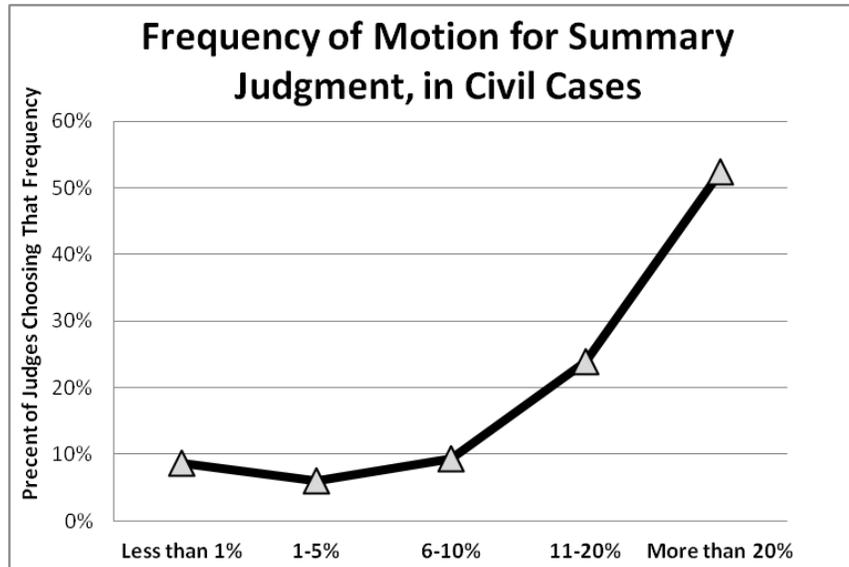
By asking judges about their experience with and handling of motions challenging reliability of expert testimony and by learning the reasons judges grant these motions, how often, and to what extent, this survey has provided an important set of new data explaining the actual practice of state court judges in courtrooms today.

<sup>114</sup> Mean = 2.43; standard deviation = 1.46.

#### *D. Judicial Handling of Summary Judgment*

Beyond the survey data on reliability motions, the survey also asked judges about the frequency of and handling of motions for summary judgment. Before I switch to comparing the data on summary judgment and reliability motions in Part III.E below, I first will discuss the responses on summary judgment alone.

The first question on summary judgment motions asked judges: “In what percentage of civil cases do you face a contested motion for summary judgment prior to trial?” In response, a majority (52%) of judges believed such motions occurred in over 20% of cases, with an additional 24% answering between 11 and 20% of cases ( $n=79, 36$ ).<sup>115</sup> Only 24% of judges believed contested summary judgment motions occurred in 10% or less of cases, with 9% answering between 6 and 10%, 6% answering between 1% and 5%, and another 9% answering less than 1% ( $n=14, 9, 13$ ).



**Figure 8: Frequency of Contested Motions for Summary Judgment in Civil Cases**

<sup>115</sup> Although 158 judges completed the survey overall, only 151 answered this question. That explains the percentage calculations here.

Just as with the questions about reliability, the survey then asked judges if they had personally ruled on a summary judgment motion and, if so, in how many cases. Of 158 respondents, 92% had ruled on a contested motion for summary judgment ( $n=145$ ). Of those who had done so, about two-thirds (66.4%) had done so in over twenty cases ( $n=95$ ). Judges selected the other options much less frequently, with 14% choosing eleven to twenty cases ( $n=20$ ), 11% choosing six to ten cases ( $n=16$ ), 7% choosing two to five cases ( $n=10$ ), and only 1% selecting one case ( $n=2$ ).

For those judges who had ruled on a summary judgment motion, the survey then asked what procedural techniques were helpful in ruling on the motion. As with the identical question with reliability motions, the possible responses included Justice Breyer's suggested methods from *Joiner* – questioning a witness from the bench, independent expert, special master, or hearings (with or without testimony presented).<sup>116</sup> Judges were again instructed to choose any factors they believed were helpful, and were also given the option of choosing an open-ended “other” response and then being permitted to enter text to explain their answer. Judicial responses to this question are recorded below in Figure 9.

---

<sup>116</sup> See *supra* note 15 and accompanying text (procedures identified in *Joiner*).

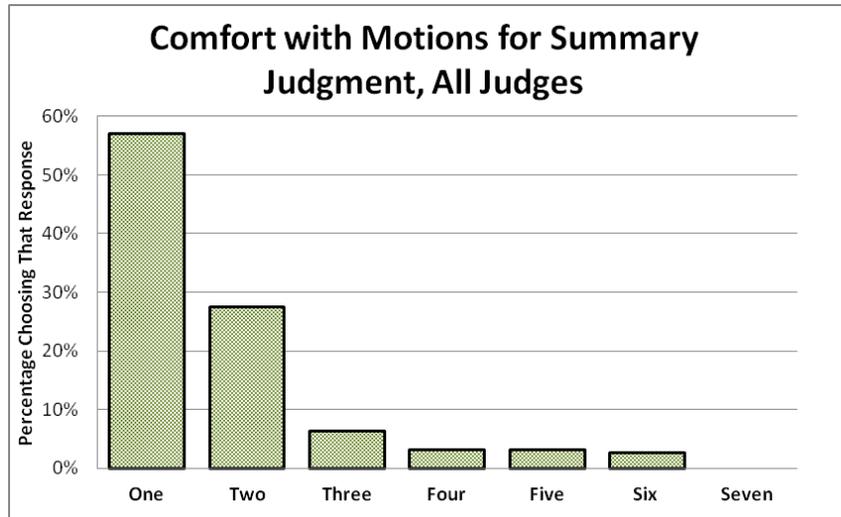
	Number Who Have Used the Technique	Percentage
Hearing With Testimony Presented	40	27.6
Hearing Without Testimony Presented	123	84.8
Questioning a Witness From the Bench	9	6.2
Independent Expert	5	3.4
Special Master	5	3.4
Other	20	13.8
	Total = 145	

**Figure 9: Procedural Methods Used by Judges to Decide Motions for Summary Judgment**

Just as with the motions for reliability, the survey asked all judges about their comfort level with handling summary judgment on a one-to-seven scale. In response, a majority of the judges - (57%) - answered the question with “1” and an additional 27.6% answered “2” ( $n=89$ , 43).<sup>117</sup> Less frequent answers included 6.4% answering “3” ( $n=10$ ), 3.2% answering both “4” and “5” ( $n=5$ , 5), and 2.6% answering “6” ( $n=4$ ). No judge in the survey responded with a “7” to this question.<sup>118</sup>

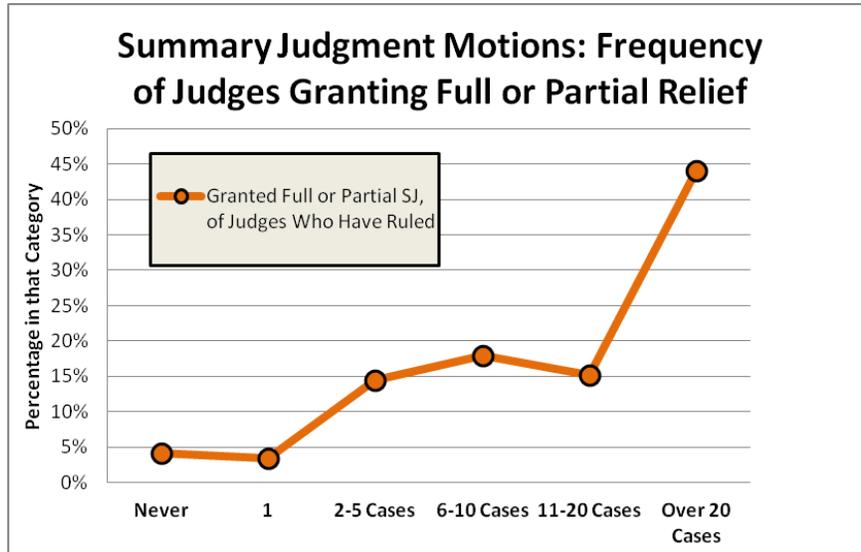
<sup>117</sup> Although 158 judges completed the survey overall, only 156 answered this question. That explains the percentage calculations here.

<sup>118</sup> Mean = 1.76; standard deviation = 1.19.



**Figure 10: Comfort Level with a Motion for Summary Judgment, All Judges**

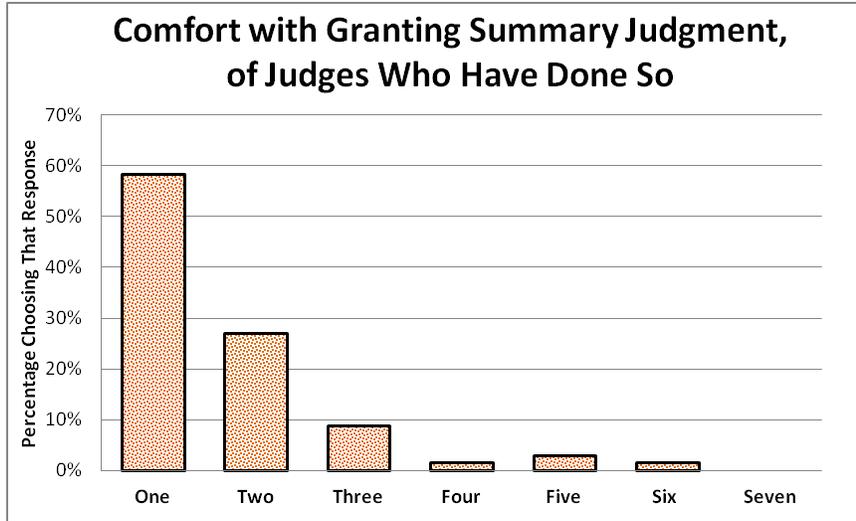
Following the initial questions on the frequency of, methods to decide, and comfort with motions for summary judgment, the next questions asked the judges who have ruled on a motion for summary judgment whether they have had granted the motion, either in full or in part. Of the 145 judges who have ruled on a motion for summary judgment, most had granted the motion in more than ten cases. Forty-four percent of judges who had granted summary judgment had done so in over twenty cases ( $n=64$ ), and an additional 15.2% had done so in eleven to twenty cases ( $n=22$ ). Of the remaining judges, 18% had granted summary judgment in six to ten cases ( $n=26$ ), 14.5% in two to five cases ( $n=21$ ), 3.4% in 1 case ( $n=5$ ), and only 4% had never granted the motion ( $n=6$ ). The variations in responses are depicted below in Figure 11.



**Figure 11: Frequency of Granting Full or Partial Relief Pursuant to a Motion for Summary Judgment, of Judges who Have Ruled on Such a Motion**

The final question asked the judges who had granted full or partial summary judgment how comfortable they were in doing so, on a one-to-seven scale. A majority of the judges (58.4%) rated their comfort level at “1” and an additional 27% answered “2.” ( $n=80, 37$ ). Only twenty judges rated their comfort level between three and six, with 8.8% selecting “3” as their response, 1.5% selecting “4”, 2.9% selecting “5” and an additional 1.5% selecting “6.” ( $n=12, 2, 4, 2$ ). No judge in the survey responded with a “7” to this question.<sup>119</sup>

<sup>119</sup> Mean = 1.68; standard deviation = 1.06.



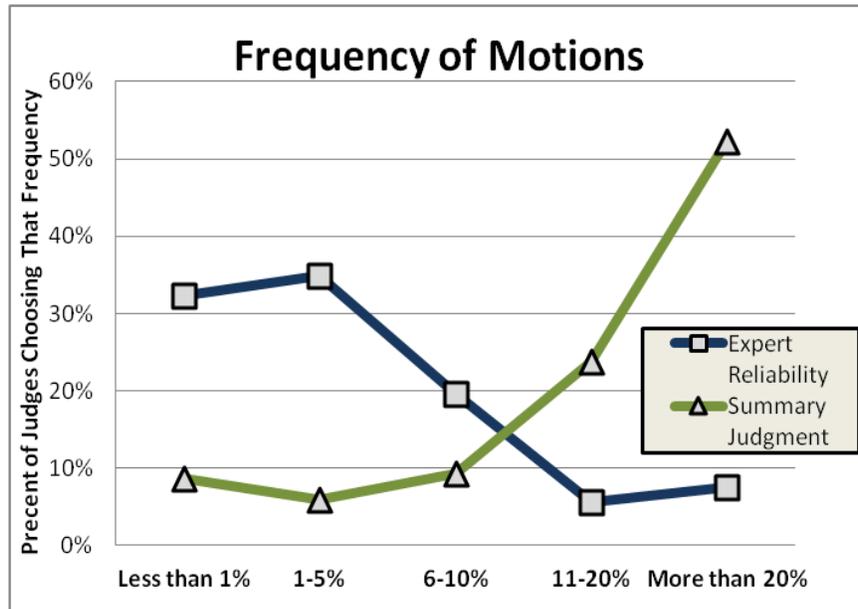
**Figure 12: Comfort Level with Granting Full or Partial Summary Judgment, Of Judges Who Have Granted the Motion**

Just as with reliability motions, the response data from judges about their experience with handling summary judgment motions, this survey has provided an additional set of interesting new data explaining the actual practices of state court judges in court.

#### *E. Comparing Reliability Motion Data to Summary Judgment Data*

The final analysis of the complete set of responses to the survey involves comparing the data on the judicial handling of motions on expert reliability to the data on summary judgment motions. In so doing, the data demonstrates several interesting new findings.

With both expert reliability and summary judgment motions, the survey asked all judges the frequency at which they face those motions. When we place the responses from Figure 1 and Figure 8, *supra*, into a single chart, the frequency of the motions compares as follows:

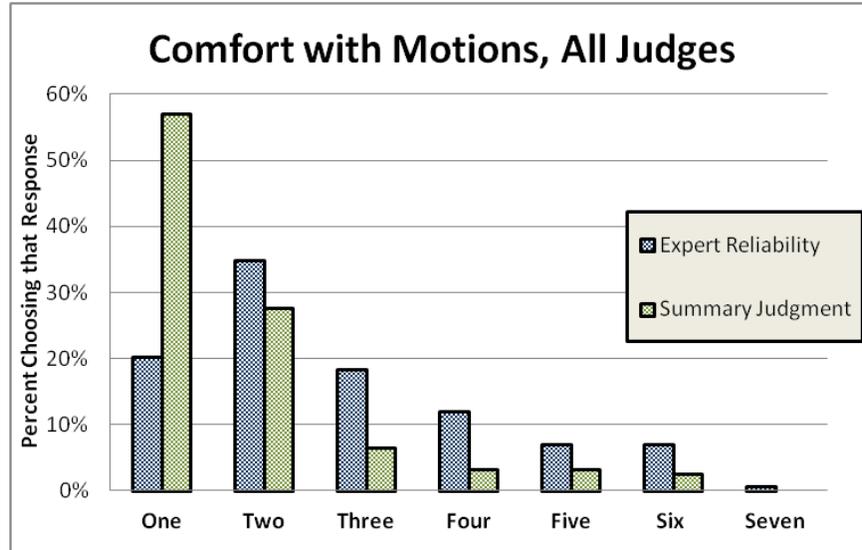


**Figure 13: Frequency of Motions Challenging Expert Reliability, in All Cases with Experts, and Contested Motions for Summary Judgment, in Civil Cases**

Clearly, a motion for summary judgment is a much more common motion for the judges to handle, with a majority handling it in more than 20% of cases, while a majority of judges see expert reliability motions in 5% or less of cases.

In addition to the question on frequency, all judges were asked about their comfort level with both motions. When we place their responses from Figure 4 and Figure 10, *supra*, into one combined chart, we can directly compare the comfort level of the judges for the two motions (Figure 14).<sup>120</sup>

<sup>120</sup> Using a *t*-test analysis, these differences are statistically significant ( $p < 0.001$ ). ALAN AGRESTI & CHRISTINE FRANKLIN, STATISTICS 388-98 (2007). I will only describe a difference in data as statistically significant if its *p*-value is less than 0.05. This corresponds to a one in twenty chance (or 5% chance) that the result is due to chance. DAVID HENSHER, JOHN M. ROSE, & WILLIAM H. GREENE, APPLIED CHOICE ANALYSIS: A PRIMER 46-47 (2005). In addition, the *p*-value of 0.05 is consistent with general practice. See, e.g., HENSHER, ROSE, & GREENE, at 46-47; AGRESTI & FRANKLIN, at 379-80.



**Figure 14: Comfort Level with a Motion Challenging Reliability of Expert Evidence and a Motion for Summary Judgment, All Judges**

The data demonstrates that judges are very comfortable in handling summary judgment and less so with the expert reliability motions. However, in both motions, most judges rate their comfort level at either “1” or “2” (55.1% for expert reliability; 84.6% for summary judgment).

In addition to the questions asked of all judges, only those judges who had faced an expert reliability motion or a motion for summary judgment were asked which techniques they used to decide the motion.<sup>121</sup> The response options remained the same for both motions so they could be compared.

<sup>121</sup> The responses on expert reliability motions are found *supra* in Figure 3 and accompanying text, and the responses on summary judgment are found *supra* in Figure 9 and accompanying text.

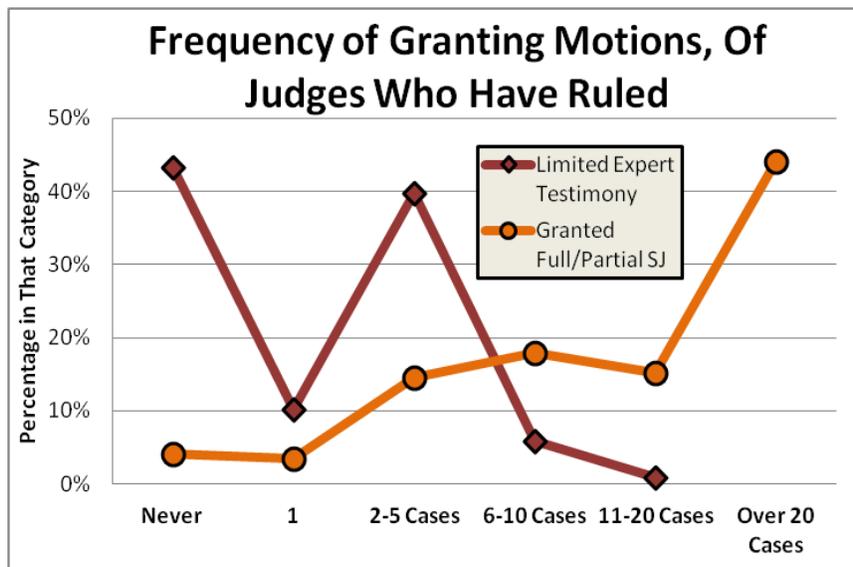
	Number Who Have Used the Technique (Expert Reliability)	Percentage	Number Who Have Used the Technique (Summary Judgment)	Percentage
Hearing With Testimony Presented	101	85.6	40	27.6
Hearing Without Testimony Presented	63	53.4	123	84.8
Questioning a Witness From the Bench	64	54.2	9	6.2
Independent Expert	9	7.6	5	3.4
Special Master	4	3.4	5	3.4
Other	10	8.4	20	13.8
	Total = 118		Total = 145	

**Figure 15: Procedural Methods Used by Judges to Decide Motions, for Judges Who Have Ruled on that Type of Motion**

For both types of motions judges commonly used hearings with or without testimony although it was more likely that testimony would be presented for reliability motions and more likely it would not for summary judgment motions. For both motions, the appointment of either a special master or independent expert was rare, with fewer than 8% of judges in the survey ever using either technique. While this data may be similar, the responses on questions from the bench were quite varied. A majority of judges who had ruled on expert reliability

had used this technique (54.2%), while few judges felt it was necessary to decide a summary judgment motion (6.2%).<sup>122</sup>

Following these questions, the survey also asked questions to the subset of judges who had ruled on and granted the motion in question. One question asked judges whether they had granted a motion challenging reliability of an expert, and if so, how often they had done so.<sup>123</sup> The survey asked an identical question to judges who had granted summary judgment.<sup>124</sup> The combined responses are displayed in Figure 16.



**Figure 16: Frequency of Limiting Testimony or Granting Full or Partial Summary Judgment Pursuant to a Motion, of Judges who Have Ruled**

As with the responses on the frequency of handling the motion, *supra* Figure 13, the data here also show higher frequencies of judges granting summary judgment than with

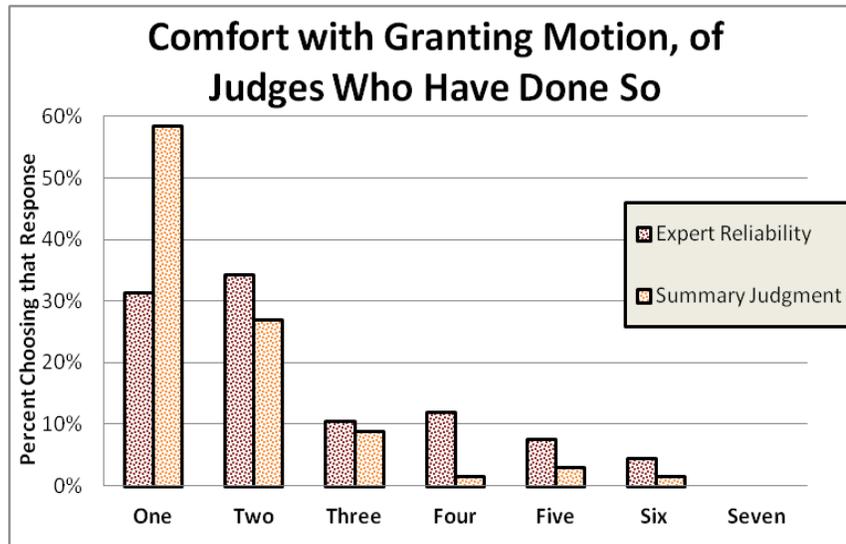
<sup>122</sup> This makes sense considering that, based on the survey responses, judges in summary judgment hearing are less likely to have testimony presented. *See* Figure 15. Without testimony, clarifying questions seem less likely to be helpful.

<sup>123</sup> *See supra* Figure 5 and accompanying text.

<sup>124</sup> *See supra* Figure 11 and accompanying text.

motions on expert reliability, although the relationship of the two is not as clear as with general frequency.

The other question asked of all judges who have granted a motion challenging reliability of an expert or a motion for summary judgment was how comfortable they felt doing so.<sup>125</sup> The combined responses are found in Figure 17.



**Figure 17: Comfort Level with Granting a Motion Challenging Reliability or a Motion Granting Summary Judgment, of Judges Who Have Granted that Motion<sup>126</sup>**

The results here again mirror the results on the question regarding general comfort with the motion. The data demonstrates that judges are very comfortable in granting summary judgment and less so with the expert reliability motions. As with general comfort, most judges rate their comfort level for

<sup>125</sup> The responses on expert reliability motions are found *supra* in Figure 7 and accompanying text, and the responses on summary judgment are found *supra* in Figure 12 and accompanying text.

<sup>126</sup> Using a *t*-test analysis, these differences are statistically significant ( $p < 0.001$ ). AGRESTI & FRANKLIN, *supra* note 120, at 388-98.

granting both motions at either “1” or “2” (68.7% for expert reliability; 85.4% for summary judgment).<sup>127</sup>

One additional analysis of the data involved isolating the subset of judges who were most comfortable handling either type of motion - answering “1” on the seven-point scale for each - and then analyzing their responses as compared to other judges. In the response group, 32 judges chose the “entirely comfortable” category for expert reliability motions. Interestingly, of those 32 judges, *every one* also rated their comfort level with a summary judgment motion in the “1” category as well.<sup>128</sup>

A comparison of the judicial responses on expert reliability motions and motions for summary judgment shows some interesting connections between the two. Clearly, there is a cohort of judges who are comfortable hearing or granting either motion, a finding that could merit further study. Differences between the motions are also apparent, however. Summary judgment motions

---

<sup>127</sup> Beyond these graphical comparisons, statistical analysis of the responses can also provide insight into the relationship of the two motions. To compare the responses of judges on their comfort with reliability and summary judgment, I calculated the effect size to further categorize the differences. Effect size calculations involve a quantification of “the degree of association between two related variables....” PAUL D. ELLIS, *THE ESSENTIAL GUIDE TO EFFECT SIZES: STATISTICAL POWER, META-ANALYSIS, AND THE INTERPRETATION OF RESEARCH RESULTS* (2010); *See generally* JACOB COHEN, *STATISTICAL POWER ANALYSIS FOR THE BEHAVIORAL SCIENCES* (2d ed. 1988). Each calculation I performed involved the comparison of a first group, such as Comfort Level with Reliability Motions, to a second, such as Comfort Level with Summary Judgment. By so doing, I could statistically analyze the relationship of Reliability Motion comfort levels to Summary Judgment comfort levels, and vice-versa. When I performed the calculations, the effect size ( $r$ ) was found to be between 0.28 and 0.36 for all of the cross-motion analyses. These  $r$ -values indicate a small-to-medium effect size under Cohen’s categorical definitions. COHEN at 79-80. *See also* ELLIS at 40-42 (describing effect size interpretation using Cohen’s descriptors). Cohen described a medium effect size as one that, in behavioral sciences, would be “perceptible to the naked eye of a reasonably sensitive observer.” COHEN at 80. Considering the two motions in question are intended to involve separate issues – admissibility and sufficiency – the finding of  $r$ -value helps to statistically quantify a closer relationship between the two.

<sup>128</sup> The finding here is similar to the subset of judges who were most comfortable granting reliability motions. Of the twenty-one judges who were entirely comfortable with limiting expert testimony due to a reliability motion, 84% also answered “1” on their comfort with granting summary judgment ( $n=16$ , of 19). The pattern does not hold for judges who were most comfortable reviewing or granting summary judgment motions. Of the 89 judges who rated “1” with summary judgment motions in general, their mean score on reliability motions was 2.22. For the 80 judges who rated a “1” on granting a motion for summary judgment, only 44% also rated a “1” with granting a reliability motion with 36% more rating as “2” ( $n=16$ , 13, of 36).

remain much more common in general frequency, and also much more common for judges to grant. It is also clear that even if many judges are comfortable with both motions, the responses support the claim that judges are slightly more comfortable with judging sufficiency of evidence under summary judgment than reliability.

#### *F. Comparing Judges Based on Home-State Admissibility Standard*

In addition to reviewing the responses of all judges, I also decided to break the judges into groups based on their home-state expert admissibility standard, to determine if there are any statistically significant differences between the judges applying a different standard.<sup>129</sup> Differences between judges applying different standards could signal the importance of the change in the substantive law of admissibility after *Daubert*, and provide a way to compare how each standard is applied.<sup>130</sup>

In analyzing the response data, many areas did not show statistically significant differences. This finding alone is interesting, since it shows that judges do not vary based on home state admissibility standard on many issues, including: comfort with handling a reliability motion, whether they have limited expert testimony based on a reliability motion (and how many times), comfort with summary judgment, number of times they have ruled on summary judgment, and number of times the judges have granted a motion for summary judgment.<sup>131</sup>

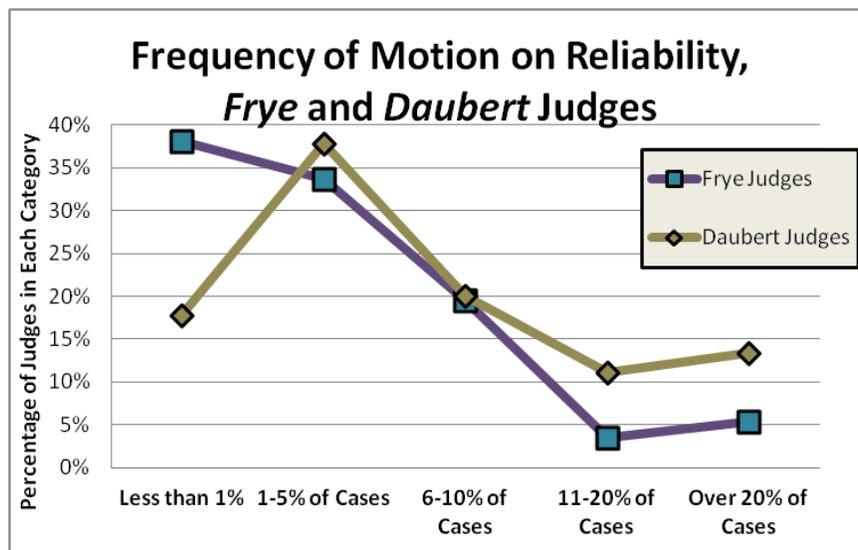
---

<sup>129</sup> For the admissibility standard of each state, see *supra* notes 90-95 and accompanying text. Analysis of all other demographic groupings is beyond the scope of this paper, and will be presented in a separate work. *Jurs*, *supra* note 89. For the list of variables, see *supra* note 108.

<sup>130</sup> In a prior study asking about use of judicial questioning and independent experts, I could find no major statistically-significant differences between judges based on home-state standards. See *Jurs*, *supra* note 88, at 65-68 (“Based on these data, the standard for admissibility of expert testimony appears to have little effect on the overall use, or frequency of use, of the advanced factfinding methods in Rule 614 and Rule 706.”). See also Gatowski, *supra* note 4, at 443 (no difference between judges of varying home-state standards on whether role as a gatekeeper is appropriate). Regarding a non-survey statistical analysis of which standard is stricter, see *Jurs & DeVito*, *supra* note 111.

<sup>131</sup> For each of these categories, the data differences, if any, lacked statistical-significance ( $p > 0.05$ ).

However, there are some interesting and significant differences between the judges based on their home-state reliability standard. The first major example deals with judicial responses to the question about the frequency of motions challenging expert reliability.<sup>132</sup> When asked what percentage of cases with experts result in a reliability motion, judges in a *Daubert* jurisdiction were significantly more likely to select a higher percentage of cases than their *Frye* counterparts.<sup>133</sup> The results are displayed graphically in Figure 18.



**Figure 18: Frequency of Motions Challenging Expert Reliability, in All Cases with Experts, By Home-State Admissibility Standard**

These are not the only statistically significant differences between judges, however. On the critical issue of which gatekeeping standard is more likely to exclude evidence, judges from *Daubert* jurisdictions and judges from *Frye* jurisdictions also differed significantly. When asked directly which standard is the

<sup>132</sup> For the response data for all judges, see *supra* Figure 1 and accompanying text.

<sup>133</sup>  $\chi^2 = 10.2$ ;  $p < 0.05$ . This result is confirmed using Fisher's exact test, where  $p$  remains less than the threshold of 0.05. See AGRESTI & FRANKLIN, *supra* note 120, at 514 (explaining that Fisher's exact test is more appropriate for small sample analysis).

stricter one for analyzing reliability, *Frye* judges were nearly evenly split: 50.4% believed that *Daubert* is the stricter standard, while 49.6% chose the *Frye* standard ( $n=57, 56$ ). On the other hand, judges who serve in a *Daubert* jurisdiction, and who therefore would have more familiarity with how that standard worked, were not evenly divided. A supermajority of 87% of those *Daubert* judges chose the *Daubert* standard as stricter, while only 13% chose *Frye* ( $n=39, 6$ ). This difference is statistically significant.<sup>134</sup>

One other statistically significant difference between judges occurred on the issue of the methodology of deciding summary judgment. When judges who have ruled on a motion for summary judgment were asked what techniques they used to rule on the motion, *Daubert* judges were evenly split on the utility of holding a hearing with testimony, with 52.4% saying they have used that technique, and 47.6% saying they had not ( $n=22, 20$ ). When the same question is asked of *Frye* judges, a vast majority of them have not used a hearing with testimony to decide on summary judgment (80.6%;  $n=83$ ) while only a few had used that technique (19.4%;  $n=20$ ).<sup>135</sup>

Independent of the frequency of reliability motions, the belief of judges on what embodies a stricter standard, and the perceived utility of hearings with testimony, all of which show statistically-significant differences between *Frye* and *Daubert* judges, there are many other factors which – while their  $p$ -value narrowly misses the threshold value of 0.05 – could later be shown to make a difference. These include: the number of cases in which judges have ruled on a reliability motion ( $p=0.061$ ), utility of the substantive factor of “technique can and has been tested” to determining reliability motions ( $p=0.069$ ), comfort level with limiting expert testimony by a reliability motion ( $p=0.057$ ), and comfort with granting full or partial summary judgment ( $p=0.053$ ).<sup>136</sup> Each of these may very well show a statistically-significant difference in a follow-up study with more respondents,

---

<sup>134</sup>  $\chi^2 = 17.71; p < 0.05$ .

<sup>135</sup>  $\chi^2 = 11.88; p < 0.05$ . To see the results for all judges, see *supra* Figure 9.

<sup>136</sup> These  $p$ -values are found using Fisher’s exact test. See *supra* note 120.

but that conclusion cannot be made based solely on the current data.<sup>137</sup>

While there are several differences between *Frye* and *Daubert* judges contained within the current data, additional study could easily result in a significantly more varied difference being found in judges' experiences with reliability and summary judgment motions.

#### IV. DISCUSSION AND SUGGESTIONS FOR FURTHER RESEARCH

Having reviewed the survey data, I will finish with some thoughts on what the survey results mean and what additional studies are implicated by this research.

##### *A. Implications of the Survey Data*

The responses of the state court judges in this survey provide important empirical findings on how often judges handle reliability motions, what tools and factors they use to decide those motions, how often they grant relief pursuant to those motions, and how comfortable they are doing so. In addition, the data offer some clues to the relationship of reliability motions to summary judgment motions, as suggested by the literature on the area.

These results update and expand upon findings of earlier research in the area of reliability motions and on the frequency, procedural methodology, and substantive factors used to assess expert admissibility. In the 1990's and early 2000's, researchers using both survey and statistical database analysis methodologies made timely and interesting findings in these areas.<sup>138</sup> With this study, we can see how trial-level judges currently handle these issues. When we look at the data, we see broad-based agreement with the prior studies. For example, prior research shows judges are reluctant to appoint an independent expert,<sup>139</sup> and the responses here confirm that result with only 7.6% of judges having used that technique to decide a reliability motion.<sup>140</sup> Another example involves substantive factors for deciding admissibility.

---

<sup>137</sup> See *infra* Part IV.B (*Suggestions for Future Research*).

<sup>138</sup> See *supra* Parts I.A.1,2, and I.B.

<sup>139</sup> See *supra* Part I.A.2 and notes 37-40 and accompanying text.

<sup>140</sup> See *supra* Figure 3.

Prior studies found that judges more often used the factors of general acceptance and peer review in deciding reliability.<sup>141</sup> The responses here are similar, with general acceptance being chosen by the respondents as the most helpful factor and peer review chosen by a large majority of the judges as well.<sup>142</sup> The responses here not only help to update, but also confirm, earlier findings in the area of reliability motions.

The study also examined whether the home-state standard of a judge affected their handling of reliability and summary judgment motions. In evaluating the data, judges who came from a *Daubert* state reported significantly more motions challenging reliability of expert testimony. Only a small percentage (17.8%) of *Daubert* judges believed reliability motions were in less than 1% of cases with experts, compared to 38.1% of *Frye* judges, while over 13% believed the motion was in over 20% of cases, compared to 5.3% of *Frye* judges. The data clearly demonstrates that *Daubert* judges perceive reliability motions as occurring more frequently; this finding is also consistent with prior studies in the area.<sup>143</sup>

In addition to addressing frequency, the data also clearly demonstrates a difference between judges on their perceptions of which reliability standard is stricter. *Frye* judges are evenly divided on which standard is stricter, with 50.4% choosing *Daubert* and 49.6% choosing *Frye*.<sup>144</sup> The *Daubert* judges are not so conflicted, as 87% of them believe the *Daubert* standard is a stricter reliability standard.<sup>145</sup> The data demonstrates that judges who have had experience applying *Daubert* are much more likely to believe it is a more restrictive standard for reliability review.<sup>146</sup>

A third finding of the study involves the relationship of reliability motions to summary judgment. In prior literature, commentators like Sanders, Vidmar, Schuck, and Finley have

---

<sup>141</sup> See *supra* Part I.B notes 46, 49, 52-53 and accompanying text.

<sup>142</sup> *Supra* Figure 2.

<sup>143</sup> See *supra* notes 20, 22, 27 and accompanying text.

<sup>144</sup> See *supra* Part III.F and text preceding note 134.

<sup>145</sup> *Id.*

<sup>146</sup> The variances between judges based on their home state standard also provide a tantalizing preview of the potential of analyzing the judges based on other demographics and differences. Potential variables are listed *supra* in note 108. The analysis of these groupings was beyond the scope of this paper, and will be presented in a separate work. Jurs, *supra* note 89.

made a connection between the *Daubert* gatekeeping function and the likelihood of summary judgment.<sup>147</sup> In so doing, they postulated, the line between admissibility (a reliability issue) blurred with a question of sufficiency (a summary judgment issue). When we look at the data, we see some preliminary indications of that theory. While the frequency of the motions varies, as does the frequency of granting the motion, some procedural connections between the two motions are clear.<sup>148</sup> The data also shows judges as very comfortable with the motions in general, but also with granting the motions.<sup>149</sup> In addition, the subset of judges who are most comfortable with reliability motions, rating their comfort level as “1” on the seven-point scale, all answered a separate question about their comfort with summary judgment as a “1” as well.<sup>150</sup> These data begin to quantify case management similarities of the two motions, and suggest further study.

In summarizing these findings, the study demonstrates the current role of the judge in gatekeeping expert evidence; how, when and why those decisions are made; and how they relate to the sufficiency standards of summary judgment. These questions involve significant public policy choices about the role of the judge, the substantive and procedural burdens of litigation, and the vanishing trial.<sup>151</sup>

### *B. Suggestions for Future Research*

Of course, this survey raises many questions, which can only be addressed through future research. Several areas of future study seem readily apparent from the results here.

Clearly an additional survey with a larger sample size would greatly assist in affirming or rebutting the results here, and clarifying results that were outside the range of statistical significance. When comparing the groups of *Daubert* and *Frye* judges, many factors were in the range of 0.05 to 0.07 for their *p*-

---

<sup>147</sup> See *supra* notes 57-65 and accompanying text.

<sup>148</sup> See *supra* Figure 15.

<sup>149</sup> See *supra* Figures 14 & 17.

<sup>150</sup> See *supra* text preceding note 121.

<sup>151</sup> See Marc Galanter, *The Vanishing Trial: An Examination of Trials and Related Matters in Federal and State Courts*, 1 J. EMPIRICAL LEGAL STUD. 459 (2004).

value. Because of this, even a small increase in respondents would likely resolve the question of whether there are real differences between these groups on how they handle reliability or summary judgment motions.<sup>152</sup>

A survey methodology, of course, is not the sole way to address these questions. In prior studies on reliability and summary judgment motions, researchers used statistical analysis of published opinions in a computer database.<sup>153</sup> Just as the question of substantive factors in assessing expert reliability can be measured with statistical analysis of databases,<sup>154</sup> that same research approach could be taken to assess procedural methodology, extents to which reliability motions are granted, or the relationship of reliability and summary judgment motions.

Beyond expanding the survey coverage or methodological choices, I have repeatedly mentioned that even with the exact database used here, additional findings may be made.<sup>155</sup> Having collected the data, I intend to analyze it separately to determine whether any background about the judge – their experience (on the bench or in practice), location of appointment, gender, experience with math and science, comfort with math and science, region, state, or any other factor – has an effect on how they manage reliability or summary judgment motions.<sup>156</sup> These analyses will explore the nature of judicial decision-making, and explore the experiences of the judge in how they assess motions before them.

Finally, the issue of the relationship of reliability motions to summary judgment is one that merits additional attention. This survey has compared the judicial handling of the two motions, their frequency, and judicial comfort with the motions.<sup>157</sup> Clearly additional empirical work could be done to explore this

---

<sup>152</sup> See *supra* Part III.F; see also notes 135-37 and accompanying text.

<sup>153</sup> See DIXON & GILL, *supra* note 4, at 15-24; Groscup et al., *supra* note 4, at 342-44; Cecil et al., *supra* note 8, at 877; Eisenberg & Lanvers, *supra* note 8, at 9-11; Cecil & Cort, *supra* note 82, at 2.

<sup>154</sup> DIXON & GILL, *supra* note 4, at 39-40 (analyzing prevalence of *Daubert* factors in database cases).

<sup>155</sup> *Supra* notes 89, 108, 129, 146.

<sup>156</sup> This will be presented in a separate work. *Jurs*, *supra* note 89.

<sup>157</sup> *Supra* Part III.F.

relationship, and more clearly affirm or refute the commentators who have connected the motions.

All of these future analyses would help clarify the findings here, or build upon the findings using additional methods, in order to further explore the issue of judicial handling of reliability and summary judgment motions.

#### CONCLUSION

When *Daubert* changed the substantive standard for expert admissibility in federal court, researchers responded with a series of studies to evaluate the basic methodology, procedures, and frequency of reliability challenges. Using varying methodologies, these researchers created basic data on how often reliability motions occurred, the procedures used to decide them, and the importance of different substantive factors of analysis. In the decade since these initial studies, there has been little additional empirical research in the area.

This survey analyzes these important issues, by providing new survey data on the frequency, procedural methodologies of, and substantive factors to decide reliability motions. In so doing, we can find how judges handle these motions in court, and whether that is consistent with prior findings.

In addition to those issues, the study also evaluated the relationship of judicial gatekeeping to summary judgment. While many commentators have connected gatekeeping on admissibility to the substantive sufficiency of evidence with summary judgment, little empirical analysis of the issue has been done. This survey asked basic questions of judges about both motions, to begin to evaluate the connections between the motions and methods judges use to decide them.

By measuring the actual practices of state court judges, this survey provides new and interesting data on *Daubert* gatekeeping, the tools used and not used by the judiciary to that end, and the connection of reliability analyses to summary judgment.

