

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Court Review: The Journal of the American
Judges Association

American Judges Association

2015

The Gatekeeper's Toolbox: A Survey on Judicial Handling of Expert-Reliability Motions

Andrew W. Jurs

Drake University Law School in Des Moines

Follow this and additional works at: <https://digitalcommons.unl.edu/ajacourtreview>

Jurs, Andrew W., "The Gatekeeper's Toolbox: A Survey on Judicial Handling of Expert-Reliability Motions" (2015). *Court Review: The Journal of the American Judges Association*. 540.
<https://digitalcommons.unl.edu/ajacourtreview/540>

This Article is brought to you for free and open access by the American Judges Association at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Court Review: The Journal of the American Judges Association by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

The Gatekeeper's Toolbox:

A Survey on Judicial Handling of Expert-Reliability Motions

Andrew W. Jurs

In the *Daubert* decision of 1993, the Supreme Court directed federal judges to screen expert evidence for reliability before admission, rejecting the “general acceptance” standard of *Frye v. United States*.¹ To ensure the appropriate level of reliability to admit expert testimony, the Court suggested a series of substantive factors for judges to analyze, such as peer review and the *Frye* “general acceptance” standard. Several years later in *General Electric v. Joiner*, Justice Breyer also suggested procedures judges could use to decide gatekeeping questions.²

In the years after *Daubert*, researchers began to evaluate how judges perform this Rule 702 reliability screening. One group of studies considered the frequency of expert-reliability challenges, finding that litigants raised reliability issues more often after *Daubert* than before. Other studies considered the methodology of expert gatekeeping by analyzing the procedures used by judges in deciding reliability questions. Others chose to focus on the substantive factors that judges consider when gatekeeping expert testimony, finding that some of the *Daubert* factors were more useful than others.

Together these studies provide some baseline data on *Daubert*'s effect, but as useful as they are, they leave significant questions unanswered. They also rely on data from the 1990s, so they may reflect courtroom standards that have changed.³

The survey discussed here is intended to address those concerns by broadly analyzing how judges perform gatekeeping. The survey was designed to answer the following questions:

- How often did judges see motions challenging the reliability of expert testimony, and how often did they grant them?
- What procedural methodologies did they use in facing reliability motions?
- What substantive factors are helpful in deciding reliability motions?
- Considering the guidance they had on how to perform gatekeeping, were they comfortable doing so?

By asking judges these questions, the survey would not only shed light on whether judges have sufficient guidance for their gatekeeping role but could also inform judges how other judges screen expert testimony in their courtrooms. Finally,

the survey would provide updated baseline data on the gatekeeping role, which could inform the debate on whether policy changes are in order.

This article will examine these issues in detail by explaining the prior research in the area, examining the methodology and results of the judicial survey, and then finishing with some thoughts on the importance of the survey's findings. By measuring the practices of judges handling reliability challenges, this survey provides significant insight into the reality of expert gatekeeping and whether the Supreme Court's guidance matches the reality in courtrooms today.

PRIOR RESEARCH ON MECHANICS AND FREQUENCY OF GATEKEEPING

In the years surrounding the *Daubert* trilogy,⁴ researchers began to explore fundamental questions about what reliability screening meant and how it was to be done. These studies were critically important in understanding the initial impact of *Daubert* by measuring the frequency of, procedures for, and substantive factors judges used to decide expert-reliability challenges.

Prior Studies Analyzing the Frequency of Daubert or Reliability Challenges

In the first decade after *Daubert*, three separate studies touched on the issue of how frequently reliability challenges occur. In only one, however, did the researchers address the absolute rate of challenges. That study, performed by Lloyd Dixon and Brian Gill and published in 2001, evaluated the rate of reliability challenges both before and after *Daubert* by examining a computerized database of reported case opinions.⁵ They found that in the four years before *Daubert*, the likelihood of a reliability issue arising was between 68 and 71% but that it had risen to between 76 and 89% in the four years after *Daubert*.⁶ The data also showed a similar increase in the likelihood of a judge finding the expert unreliable. They concluded that these increases suggest that under *Daubert*, reliability standards had tightened, which encouraged litigants to file more challenges to opponents' experts.⁷

This article is an abridged and revised version of a report of the author's survey originally published as *Gatekeeper with a Gavel: A Survey on Judicial Management of Challenges to Expert Reliability and Their Relationship to Summary Judgment*, 83 Miss. L.J. 325 (2014).

Footnotes

1. *Daubert v. Merrell Dow Pharm.*, 509 U.S. 579, 593-95 (1993); see *Frye v. United States*, 293 F.1013 (D.C. Cir. 1923).
2. *Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 147 (1997) (Breyer, J., concurring).
3. One major reason to believe courtroom standards may have changed since these prior studies is that many states, evaluating

their state evidentiary standards, switched from *Frye* to *Daubert* in the mid to late 1990s. So during the data-collection periods of several of the prior studies, some states adhered to *Frye* but have changed to *Daubert* since. For more on this issue, see *infra* Part II.d and text accompanying notes 22-25.

4. The trilogy comprises *Kumho Tire Co. v. Carmichael*, 526 U.S. 137 (1999), *Joiner*, 522 U.S. at 136, and *Daubert*, 509 U.S. at 579.
5. 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).
6. *Id.* at Table 4.1.
7. *Id.* at 29.

The next year saw another database study of *Daubert* by a group of researchers led by Jennifer Groscup.⁸ As with Dixon & Gill, their study evaluated reported case opinions from a computerized database in the years surrounding *Daubert*, except in this study, the researchers were examining criminal rather than civil cases. Yet the study only measured the frequency of admitting experts at the trial and appellate levels rather than the absolute rate of challenges. They found that more than 74% of experts had been admitted at trial and that the rate on appeal remained above 69%, although the admission rate varied dramatically between prosecution and defense experts.⁹

Beyond those two database studies, Carol Krafska and her colleagues published a survey in 2002 also touching on the issue of frequency of reliability challenges.¹⁰ Relying on surveys of state court judges from 1991 and 1998, the Krafska study found that litigants raised reliability challenges in a motion *in limine* at much higher rates after *Daubert* than before.¹¹ The authors noted, however, that their survey did not and could not address the general frequency of reliability challenges.¹²

Prior Studies Analyzing the Procedures Judges Use to Decide Reliability Motions

As with the studies addressing frequency of reliability challenges, the most recent data on the procedural methodology of gatekeeping had also been collected in the late 1990s. The Krafska study, relying on surveys of federal judges performed in 1991 and 1998, specifically asked judges what procedures they used in all cases involving experts as well as what procedures they used in complex expert cases. Her survey found that some methods, like pretrial conferences or hearings on reliability, were commonly used in all cases with experts, while others, such as independent experts or special masters, were reserved for cases with complex expert issues.¹³ The Krafska study provided a comprehensive snapshot of the methodologies used by federal judges to resolve *Daubert* challenges in the years surrounding *Daubert*.

Shirley Dobbin and her colleagues performed a survey of both federal and state court judges in 1999, which also asked about strategies for handling expert evidence.¹⁴ As with Krafska's study, the researchers asked which methodologies were used in all cases with experts and which were only for more complex or difficult cases. The Dobbin study found that state court judges were less likely than their federal counterparts to ask questions from the bench under Rule 614 or to ask the parties for instruction or education on the area of expertise.¹⁵ On the other hand, state court judges were more likely than federal judges to use an independent expert under Rule 706.¹⁶

Prior Studies Analyzing the Substantive Factors Judges Consider on Reliability

Finally, several of these studies also touched upon the substantive factors judges used to determine reliability motions, both by database analysis and by survey methodologies. Two studies—Dixon & Gill's and Groscup's—analyzed how often certain terms, the “*Daubert* factors” from the original *Daubert* opinion, appeared in reported cases during the 1990s. Dixon & Gill analyzed these terms in reported decisions of civil cases and found that judges most commonly analyzed general acceptance and peer review in making reliability choices.¹⁷ Groscup's findings are quite similar in this area. Her study found that when examining the appearance of the “*Daubert* factors” in criminal cases, judges were most likely to analyze general acceptance and peer review, although she is careful to note that all of these terms appear less frequently than more general terms such as relevance, reliability, or qualifications.¹⁸

In addition to the computerized database analysis, one published study addressed the substantive factors in gatekeeping using a survey methodology. Relying on surveys collected in 1999, Sophia Gatowski and her colleagues published a study in 2001 regarding how judges decide gatekeeping motions.¹⁹ As with the database analysis, Gatowski et al. found that judges were most likely to consider general acceptance and peer review in assessing reliability challenges, with 93% and 92% answering each factor was useful, respectively.²⁰

Conclusion Regarding Prior Research into Judicial Gatekeeping

In each of these areas—frequency, procedures, and substance—research in the first decade after *Daubert* established some baseline data about judicial gatekeeping. Why, then, is further research needed? Several important considerations lead to the conclusion that updating prior work in this area is necessary.

First, the surveys and gatekeeping databases for these studies are from the 1990s, which partially explains their import in the years right after *Daubert*. Yet the date of data collection suffers from a major weakness: The three decisions of the “*Daubert* trilogy” had not all been finalized when the data was collected. Since the last decision—*Kumho Tire* in 1999—expanded the gatekeeping role to non-scientific technical expertise, judges may have had to rethink their gatekeeping approaches after that decision.²¹

Second, when the survey involves state court judges, the *Daubert* case may not be the correct starting point anyway. For federal judges, *Daubert* had rejected the *Frye* standard for gate-

8. 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).

9. *Id.* at 345-46.

10. Carol Krafska et al., *Judge and Attorney Experiences, Practices, and Concerns Regarding Expert Testimony in Federal Civil Trials*, 8 PSYCHOL. PUB. POL'Y & L. 309 (2002).

11. *Id.* at 321.

12. *Id.*

13. *Id.* at Table 5.

14. Shirley A. Dobbin et al., *Federal and State Trial Judges on the Prof-*

fer and Presentation of Expert Evidence, 28 JUST. SYS. J. 1 (2007).

15. *Id.* at Tables 2 & 3.

16. *Id.*

17. DIXON & GILL, *supra* note 5, at 39.

18. Groscup et al., *supra* note 8, at Table 5.

19. Sophia I. 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).

20. *Id.* at 447.

21. 526 U.S. at 147.

keeping in favor of a reliability approach. But in the states, judges had to wait until their own state supreme courts decided whether *Frye's* “general acceptance” test or *Daubert's* reliability standard would remain the state gatekeeping standard. In the 1990s, many states adopted a *Daubert*-like reliability standard,²² but others did not, choosing to remain with *Frye*.²³ Since the survey data had been collected during this transition period, a follow-up survey could capture the more stable current environment.

Third, while the prior studies evaluated many aspects of gatekeeping, they didn't and couldn't cover them all. The survey could therefore be designed to update prior survey findings and expand into additional areas not covered before.

Finally, research since *Daubert* has suggested that judges might be having difficulty with the gatekeeping role, particularly with new or cutting-edge science.²⁴ Whether or not this is true, an accounting of the methodology of gatekeeping in modern courtrooms could establish which tools are in use and which are not, informing the policy debate surrounding *Daubert*.

SURVEY DESIGN

So to establish baseline data about gatekeeping methodologies and update prior work in the area, I began to design a survey tool as well as think about which judges should answer it. The survey required three main design choices, as follows:

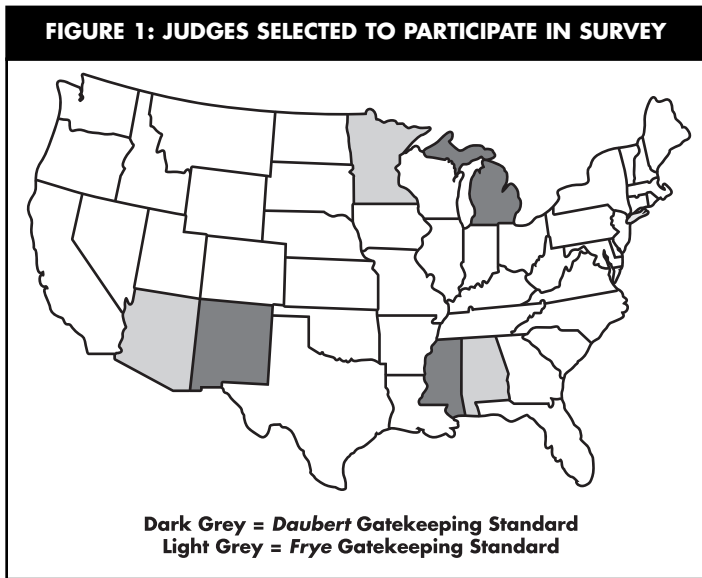
- Which judges should be involved?
- Which states should the judges come from? and
- What specific questions should be asked regarding gatekeeping?

With detailed front-side planning, I could maximize the scope and impact of the findings by making deliberate and judicious choices about who to involve and what to ask them.

Selecting the judges was the first issue to finalize, and so it became necessary to decide which judges would be “in” and which would be “out” and to have a principled reason for this distinction. After considering the alternatives, I ultimately decided to involve only the state court judges who sat on the bench of the “highest” trial court in their state. These judges are often (but not always) the most experienced, so they would be likely to be familiar with the procedures in the study. Even more importantly, they would also have the jurisdiction to hear the most complex civil cases, in which expert-reliability challenges would likely arise.

Once the judges were selected, the next step was to decide where to find them. The selection of states offered an opportunity

to perform a natural experiment. By selecting states from different regions of the United States as well as states that had different admissibility standards for expert gatekeeping, the study could examine whether regional culture or the home-state reliability standard has an effect on the way judges analyze reliability. To evaluate those considerations, the study would involve several different regions of the U.S., and in each region there would be one state with *Frye* as the home-state standard and one with *Daubert* as the standard. Furthermore, the underlying rules of civil procedure would have to be as broadly compatible as possible.²⁵ Considering these factors, the study incorporates three regions of the U.S.—West, Midwest, and South, as follows:



Limiting participation to these six states and only those judges previously mentioned, the survey began with a total of 996 eligible participants. Each judge received a mailed letter asking him or her to participate in an online survey and an additional reminder after several months.

Finally, to broadly evaluate the methodology of gatekeeping, I needed to decide specifically what questions to ask. To establish the frequency with which judges handle these motions, the survey asked them how often they see reliability motions, how often they rule on them, and how often they grant them. To see the methodologies of handling reliability motions, the survey asked the judges about the procedures they used to decide them as well as which substantive factors

22. For a list of states that have adopted a *Daubert*-type analysis, see Alice B. Lustre, *Post-Daubert Standards for Admissibility of Scientific and Other Expert Evidence in State Courts*, 90 A.L.R. 5th 453 (2001). See also DAVID H. KAYE, DAVID E. BERNSTEIN & JENNIFER L. MNOOKIN, *THE NEW WIGMORE: EXPERT EVIDENCE* § 6.4.2.a n.16 (2d ed. 2004).

23. For a list of *Frye* states, see KAYE ET AL., *supra* note 22, at § 6.4.2.a n.17; Lustre, *supra* note 22, at 453.

24. See, e.g., Andrew W. Jurs, *Balancing Legal Process with Scientific Expertise: A Comparative Assessment of Expert Witness Methodology in Five Nations, and Suggestions for Reform of Post-Daubert U.S.*

Reliability Determinations, 95 MARQ. L. REV. 1329 (2012); Andrew W. Jurs, *Judicial Analysis of Complex & Cutting-Edge Science in the Daubert Era: Epidemiologic Risk Assessment as Test Case for Reform Strategies*, 42 CONN. L. REV. 49 (2009); Gatowski et al., *supra* note 19, at 442 (judges split on whether they believe they have the necessary background to handle scientific evidence in their courtroom).

25. For a detailed explanation of this consideration, please see the complete study, *Gatekeeper with a Gavel: A Survey on Judicial Management of Challenges to Expert Reliability and Their Relationship to Summary Judgment*, 83 MISS. L.J. 325, 341 n. 101 (2014).

were most helpful in guiding those decisions. Judges were also asked how comfortable they were with reliability motions as a way to distill their thoughts about the motions to one basic concept. Finally, to have a point of comparison, I also decided to ask the judges about a different type of motion—the summary-judgment motion.

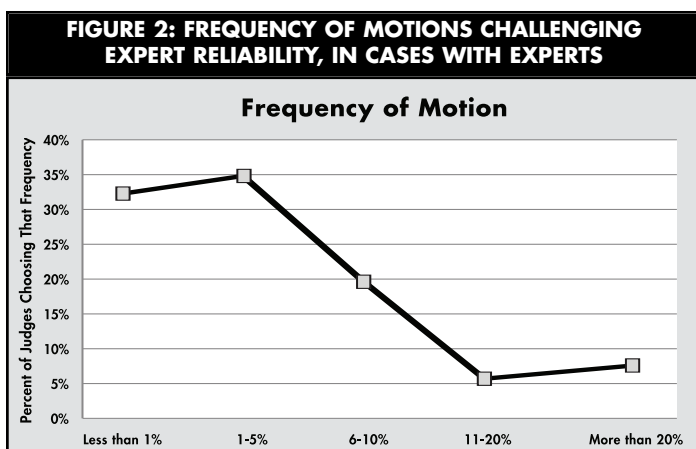
Once the design was complete, the survey went to the pre-selected eligible judges, resulting in 158 complete responses to all questions.²⁶ These responses form a unique dataset, providing a great deal of information about judicial gatekeeping in courtrooms today.

RESULTS

In providing the results, I will start by exploring the judicial responses to questions about handling expert-reliability motions. I will then compare those responses to the responses on summary-judgment motions and compare responses between different groups of judges.

Expert-Reliability Motions

The survey first addressed the frequency of expert-reliability motions. I asked the judges how often they see a motion challenging the reliability of an expert in cases with experts. The judges overwhelmingly believed this was an uncommon occurrence, with 32% answering that it occurred in less than 1% of cases with experts and an additional 35% answering that it occurred between 1% and 5% of the time. The responses to this question can be seen in Figure 2.



I next asked the judges about how they decide reliability motions, including questions on the substantive factors that guide those decisions and the procedures they have used in making those choices. On the issue of substantive factors, I provided judges with a list of the “*Daubert* factors” and asked them to choose any factors that they believed are helpful in deciding reliability motions. The responses indicate that general acceptance is the most helpful substantive factor in deciding reliability motions, with over 96% of judges selecting it. On the other hand, judges were least likely to choose error rate as a helpful factor for them, with only 70% choosing it from the list. The responses to this question appear below, as Figure 3.

FIGURE 3: SUBSTANTIVE FACTORS JUDGES BELIEVE ARE HELPFUL IN DECIDING RELIABILITY MOTIONS

	NUMBER WHO BELIEVE THE FACTOR IS HELPFUL	PERCENTAGE
Technique Can and Has Been Tested	143	90.5
Subjected to Peer Review/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	

To find which procedures judges use to decide reliability motions, I decided to ask only those judges who had ruled on this type of motion before so that the responses would reflect those procedures actually used, not those judges might use. As with the question about substantive factors, I provided judges with a list of potential procedures and asked them to select any option that they used. As with the previous question, these factors were not selected out of thin air but instead were those procedures that Justice Breyer suggested for reliability analysis in his concurrence in *General Electric v. Joiner*.²⁷ In reviewing responses, an overwhelming majority of judges had used a hearing with testimony presented to decide a reliability motion, while a slim majority of judges used a hearing without testimony or questioning a witness from the bench. On the other hand, it was very rare for a judge to use a special master or independent expert to decide a reliability motion. These responses are displayed in Figure 4.

FIGURE 4: PROCEDURAL METHODS USED BY JUDGES TO DECIDE RELIABILITY MOTIONS

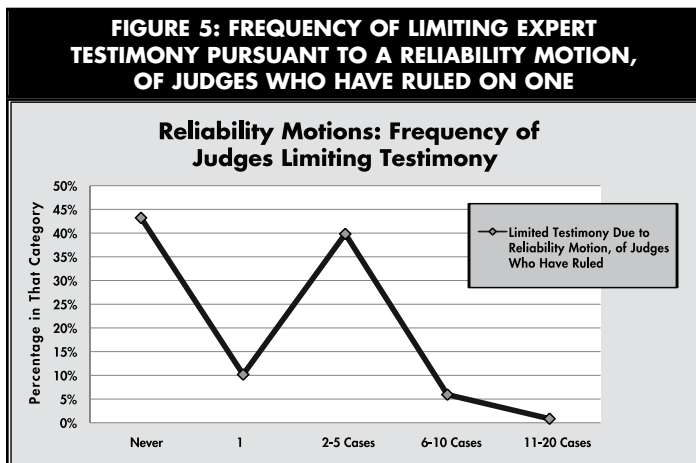
	NUMBER WHO HAVE USED THE TECHNIQUE	PERCENTAGE
Hearing With Testimony Presented	101	85.6
Hearing Without Testimony Presented	63	53.4
Questioning a Witness from the Bench	64	54.2
Independent Expert	9	7.6
Special Master	4	3.4
Other	10	8.4
	Total = 118	

So having found what factors judges consider and what procedures judges use to decide reliability motions, I then asked the judges how often they granted a motion to limit expert testimony. As with the question regarding procedures, this question was asked only of those judges who had ever ruled on a reliability motion. The responses indicate that limiting testi-

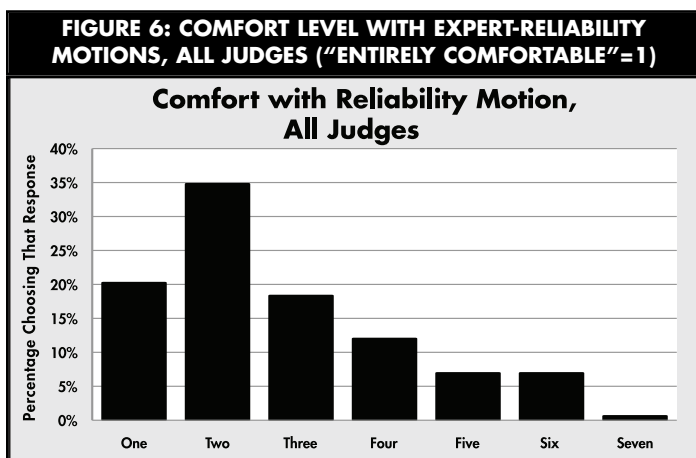
26. Thank you to all judges who took the time to participate.

27. 522 U.S. at 147 (Breyer, J., concurring).

mony is quite rare. Of the judges who had ruled on a reliability motion, 43% had never granted that motion and limited expert testimony. For the remainder who had limited expert testimony, almost all of them had done so five or fewer times, with only 7% of the judges who had ruled on this motion having limited testimony six or more times. These responses appear in Figure 5 below.



Finally, I wanted to ask judges whether, considering the procedures at their disposal and the factors they must consider, they are comfortable with expert-reliability motions. Each of these questions involved quantifying their comfort level with the motion on a seven-point scale, with one representing “entirely comfortable” and seven representing “not comfortable.” I first asked all judges how comfortable they were with expert-reliability motions, and in response, judges’ answers were quite varied. Only 20% of judges indicated they were “entirely comfortable” with the motion (category one), with an additional 35% of judges answering they were mostly comfortable (category two). On the other hand, 45% of judges responded to this question by choosing categories three to seven. These responses appear below, as Figure 6.



28. In addition to allowing a comparison of the motions, the other reason for this choice was a lingering question in the literature about the relationship between reliability challenges and suffi-

I also asked a second question to those judges who had granted a motion challenging expert reliability—about their comfort level in limiting expert testimony. As compared to the group of all judges, these responses were slightly more favorable, with 65% of judges answering in categories one or two, as opposed to 55% of all judges. In addition, no judge answered “not comfortable” in response to this question, so the remaining 35% all appeared in categories three through six. The comfort-level responses to this question appear as Figure 7.



The combined responses to the questions about expert-reliability motions contained in Figures 2 through 7 establish some baseline data about the frequency of expert-reliability challenges, how judges decide them, and how often they are granted. But I decided, when designing the study, that the results would not be as useful as possible when standing alone. Instead, I also asked about a different type of motion—summary judgment—as a point of comparison.²⁸ With the judges’ responses to these questions, I could compare the motions in many areas.

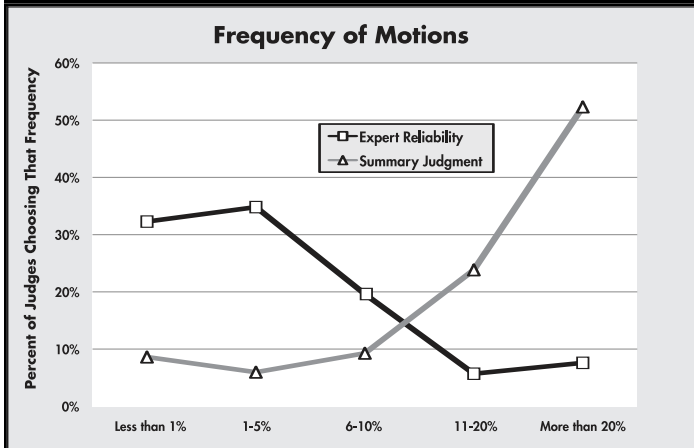
Comparing Expert-Reliability to Summary-Judgment Motions

In examining the frequency of the two types of motions, one can immediately determine that summary judgment is a much more common issue for judges to handle. When I asked the judges in what percentage of civil cases they saw a contested motion for summary judgment, a majority (52%) indicated this happened in over 20% of all civil cases, and an additional 24% indicated between 11 and 20% of all civil cases. Unquestionably, the pattern for summary judgment is different than expert reliability, as displayed in Figure 8.

Next I decided to ask judges about the methodology of deciding summary-judgment motions, to compare those responses to expert-reliability motions. The responses indicate significant differences in how judges handle these motions: for summary judgment, only 28% of judges use a hearing with testimony (86% for reliability) and 6.2% question a witness from the bench (54% for reliability). On the other hand, independent experts and special masters remain rare in both instances. The responses can be compared using the chart in Figure 9.

ciency challenges (summary judgment). For a detailed explanation of this issue, please see the complete study, *supra* note 25, at 335-39.

FIGURE 8: FREQUENCY OF EXPERT-RELIABILITY AND SUMMARY-JUDGMENT MOTIONS



Responses about the judges' comfort level with both motions provides a final point of comparison between the motions. When asked about expert-reliability motions, only 55% of judges answered they were entirely or mostly comfortable with the motion (categories one and two).²⁹ Judges were much more comfortable with summary-judgment motions. When I asked all judges to rate their comfort level with these motions on the seven-point scale, 57% answered they were entirely comfortable, and an additional 28% answered mostly comfortable, for a total of 85% in categories one and two. The judges' responses to the question about their comfort level with both motions appear graphically as Figure 11.

FIGURE 9: PROCEDURAL METHODS USED BY JUDGES TO DECIDE MOTIONS

	NUMBER WHO HAVE USED THE TECHNIQUE (EXPERT RELIABILITY)	PERCENTAGE	NUMBER WHO HAVE USED THE TECHNIQUE (SUMMARY JUDGMENT)	PERCENTAGE
Hearing With Testimony	101	85.6	40	27.6
Hearing Without Testimony	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	

The next comparison deals with the likelihood of a judge granting the motion. I asked judges how often they have granted a contested motion for summary judgment in whole or in part. As with reliability motions, this question was only for those judges who have ruled on this type of motion. In response, over 44% of judges indicated they had granted summary judgment in over 20 cases, while only a very few judges (4%) had ruled on such a motion but never granted it. When examined graphically, the distinction in the responses between these motions is clear:

FIGURE 10: FREQUENCY OF GRANTING EITHER EXPERT-RELIABILITY OR SUMMARY-JUDGMENT MOTION

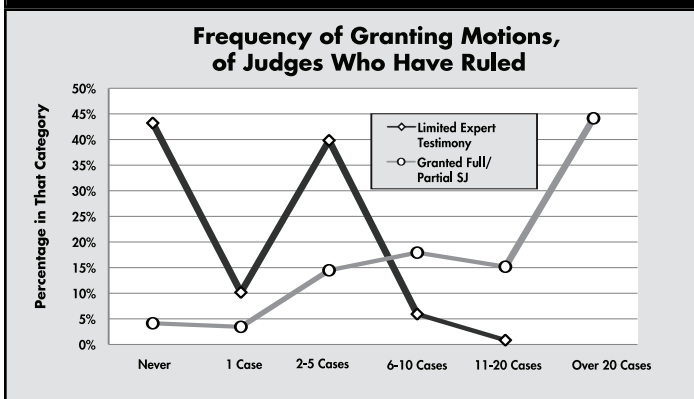
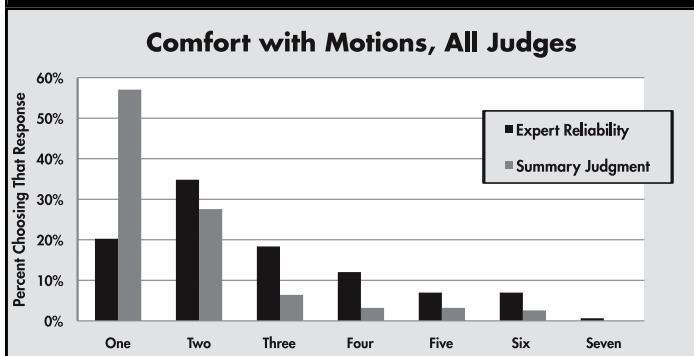


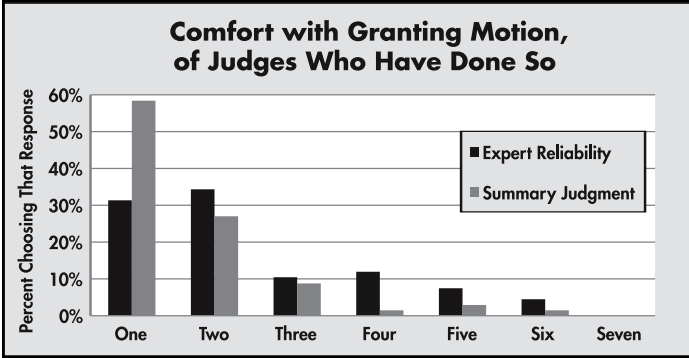
FIGURE 11: COMFORT LEVEL WITH EACH MOTION, ALL JUDGES ("ENTIRELY COMFORTABLE"=1)



The same pattern is true for the comfort level with granting the motions. For expert-reliability motions, 65% of judges answered they were entirely or mostly comfortable with granting the motion. As with summary judgment generally, the comfort level for granting summary judgment significantly exceeded the reliability number. Just as with the general comfort with summary judgment, 85% of the judges responded they were entirely or mostly comfortable with granting summary judgment, in categories one or two. The judges' responses to these questions appear below in Figure 12.

29. *Supra* Figure 6.

FIGURE 12: COMFORT LEVEL WITH GRANTING EACH MOTION, OF JUDGES WHO HAVE DONE SO



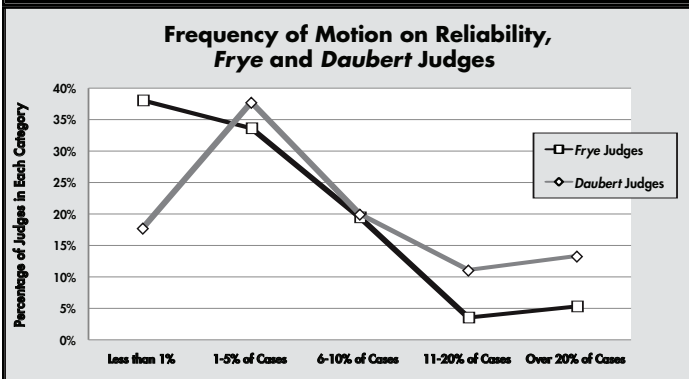
Comparison of Responses by Judges’ Backgrounds

When I designed the survey, I had been very careful in choosing which states would participate so that I could perform several natural experiments. In selecting the six states I used, I could split the responses I received into different groups and then re-examine them based on state, region, and also the home-state expert admissibility standard. I also asked judges about their backgrounds, with questions about their years of experience on the bench, years in practice, and training or comfort level with math and science.

When I evaluated these different groupings, what surprised me most was how few differences existed between categories of judges. While a smattering of differences arose, the main category where judges answered questions differently involved the home-state gatekeeping standard. These differences arose in response to two questions.

The first deals with the frequency of facing reliability motions, a question discussed above and reported in Figure 2. When the complete set of judicial responses was split, however, between judges from *Daubert* states and judges from *Frye* states, the result does show a clear difference: *Daubert* judges face more reliability motions. In Figure 13, the responses from both groups are reported, and the responses indicate *Daubert* judges are more likely to believe reliability motions happen in 11% or more of their cases and less likely to believe that expert-reliability motions occur in a very small percentage (less than 1%) of their cases.

FIGURE 13: FREQUENCY OF EXPERT-RELIABILITY MOTIONS, BY HOME-STATE GATEKEEPING STANDARD



The frequency question provided one difference between *Frye* and *Daubert* judges, and the other difference occurred in response to a final question in which I asked directly: “Which standard is the stricter one for reliability of scientific evidence—*Frye* or *Daubert*?” When I asked *Frye* judges this question, they were evenly split, with 50.4% answering *Daubert* and 49.6% answering *Frye*. The *Daubert* judges’ responses were quite different, however. An overwhelming majority of those judges—87%—believed the *Daubert* standard was stricter than *Frye*. These responses appear in Figure 14.

FIGURE 14: FRYE OR DAUBERT AS A STRICTER TEST, BY HOME-STATE GATEKEEPING STANDARD

		HOME-STATE GATEKEEPING STANDARD	
		Frye	Daubert
WHICH IS STRICTER?	Frye	50.4%	87%
	Daubert	49.6%	13%

DISCUSSION AND CONCLUSION

The survey responses discussed above provide baseline data on the frequency and handling of expert-reliability motions in courtrooms today as well as insight into judicial attitudes about such motions. Having reviewed the response data, I would like to highlight those results of the study that were most interesting.

The intent of the study was to both update and expand prior research in the area of expert gatekeeping. The data do seem remarkably consistent with the prior studies in the area to the extent they have measured these issues. For example, all three research studies from the 1990s that measured substantive factors for gatekeeping found that the most important substantive factors in deciding an expert-reliability motion were general acceptance and peer review.³⁰ Judges responding to this survey found general acceptance the most helpful substantive factor and peer review the third most helpful (behind testing). Beyond substantive factors, the results also are broadly similar to prior studies on procedures too. Prior studies showed judges were unlikely to use an independent expert in their courtrooms, and this survey confirms that finding, with only 7.6% of judges using that technique for a reliability motion.

Beyond confirming prior research, however, the study here does branch into new areas. It provides baseline data on the frequency of expert-reliability motions and also on how frequently they are granted. The study also had judges distill their opinions about expert-reliability motions into one basic concept—“comfort level”—and it shows judges are somewhat comfortable with reliability motions but not “entirely comfortable” either.

The survey is also useful in being able to compare responses about reliability motions, like the “comfort level” answers, with another common type of motion, summary judgment. By comparing the two, the survey shows us that judges are significantly less comfortable with motions about reliability than

30. DIXON & GILL, *supra* note 5, at 39; Groscup et al., *supra* note 8, at Table 5; Gatowski et al., *supra* note 19, at 445-47.

with summary judgment. The responses also show judges face many more motions for summary judgment in their courtrooms than reliability challenges and are more likely to have granted a motion for summary judgment as well.

Finally, the survey results can be split into subgroups, and by comparing the groups, we can see differences in the answers between home-state *Daubert* judges and home-state *Frye* judges. The judges from *Daubert* states reported a higher frequency of expert-reliability challenges than their *Frye* counterparts. When asked to compare the two standards, *Daubert* judges believed their own standard was stricter, while *Frye* judges were evenly split between the two standards. These responses indicate that if a judge has used the *Daubert* standard, that judge is more likely to believe it is stricter.³¹

By asking judges about their handling of expert-reliability motions, this study provides baseline data about how judges decide these motions and how often they see them; in doing so, it informs the policy debate about whether the current tools at their disposal are appropriate to the task.



Andrew Jurs is an Associate Professor of Law at Drake University Law School in Des Moines, Iowa. His research focuses on judicial handling of complex scientific evidence, including the empirical analysis of the methodology of gatekeeping and its effect within the judicial system. He teaches Evidence and Scientific & Expert Evidence in his field of research, as well as Criminal Law and Criminal Procedure, which build upon his experience as a Deputy District Attorney in Colorado. Professor Jurs received his law degree from the University of California and his undergraduate degree from Stanford University. E-mail: andrew.jurs@drake.edu.

31. In addition to studying this question with surveys, a co-author and I have been studying which standard is stricter using a statistical approach. In those studies, we found that—using a database of millions of real cases—civil litigants act in ways demonstrating that *Daubert* is a stricter standard. Andrew W. Jurs & Scott DeVito,

Et Tu, Plaintiffs? An Empirical Analysis of Daubert's Effect on Plaintiffs, and Why Gatekeeping Standards Matter (a Lot), 66 ARK. L. REV. 975 (2013); 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).