

The Heuristic Tendency in Federal District Court Rulings

Andrew B. Dzeguze
Visiting Assistant Professor
Whittier College Department of Political Science

Submitted to the 2018 Western Political Science Association Annual Meeting

San Francisco, California

March 28-31, 2018

Abstract

Over the last several decades, multiple competing schools of thought have emerged regarding what impacts judicial decision-making. An emerging body of experimental data has suggested that judges are influenced by heuristics and other cognitive shortcuts at the initial decision making stage which could lead to sub-optimal decisions. These articles have suggested, however, that the process of authoring opinions might limit or counter the impact of such tendencies.

The present paper is an outgrowth of an assessment of Federal district court decisions on the admissibility of expert witnesses. Employing thematic analysis of all cases involving a substantive analysis of this issue from 2010-2015 in nine district courts within three different circuit courts of appeal, it appears that judges have adopted several practices in crafting their written product consistent with minimizing the cognitive burden of decision-making. These include a choice architecture that encourages the admission of expert witnesses wherever possible, substitution of prevailing legal norms and industry standards for independent assessment of the evidence, and the use of a radically simplified legal framework in most cases that emphasizes tasks that judges are more familiar with such as relevance determinations over complex methodology questions. These findings suggest that the cognitive shortcuts undertaken in the process of drafting rulings may reinforce rather than ameliorate sub-optimal outcomes in many circumstances.

Table of Contents

Abstract	i
Table of Contents	ii
I. Introduction and Overview	1
II. Data And Methods	8
III. Results and Discussion	12
A. The Admissibility Heuristic - Excluding an Expert is an Unusual Step.....	14
B. The Acceptance Heuristic - Judges Substitute Third Party Standards and Prior Cases for Their own Assessment of Methods and Particular Experts.....	23
C. The Framework Heuristic - Judges' Rulings Demonstrate Significant Cognitive Simplification in Terms of Structure and Substance.....	40
IV. Conclusion	47
References	51

I. Introduction and Overview

As noted by both Rachlinski and Wistrich (2017) and Epstein (2016), empirical studies of judicial decision making have proliferated over the last several decades. Political influences, behavioral traits, panel effects, strategic and institutional considerations have all been shown to be correlated with outcomes in particular contexts. More recently, experimental work spearheaded by Guthrie, Rachlinsky and Wistrich (2007) has shown that judges are prone to many of the same cognitive shortcuts as the population at large. These cognitive shortcuts or heuristics are thought to be an outgrowth of the human tendency to minimization of cognitive effort. (Kahneman 2011). Although these shortcuts are useful in many situations, they can lead to sub-optimal outcomes as well, especially when the shortcut is predicated on faulty assumptions or arbitrary factors.

To date only limited analysis exists of how the human tendency to seek cognitive shortcuts manifests in actual judicial conduct. Beyond the experimental setting, much of the discussion has been theoretical, such as Baum's (2010) hypothesis that heuristics are at work as part of his overall call to broaden the motivations considered when analyzing judicial conduct. Similarly, Epstein and Knight (2013) suggest a role for heuristics as part of a revised approach to assessing strategic and institutional influences on judicial decision making but go no further.

Given the findings of Guthrie, et al. (2007), decades of research in general cognitive psychology and the cognitive demands of legal process, it would be somewhat surprising if courts at all levels did not adopt a heuristic process in addressing their duties. The limited work in the field suggests a role for them beyond experiment settings. For example Beebe's (2006) study of trademark infringement cases suggests that judges in different circuits use certain considerations out of a non-exclusive set of factors as near-dispositive heuristics for ultimate

decisions. Similarly, Bainbridge and Gulati (2002) postulate that doctrines justifying termination of securities law disputes serve as heuristic means of avoiding the underlying complexities of this area of law.

At the same time, there has been a suggestion that heuristics may not have much influence on judicial decisions that are embedded in written opinions. Guthrie, et al. (2007) posited that the writing process might operate as a time of reflection and de-biasing of many heuristics. This is tied to the core concept that heuristics emerge as a result of the human mind's differentiated thought processes between reflexive (or System I) and reflective (or System II) (Kahneman 2011). Writing might force the judge to be more reflective and consider if their initial impression of the issue is appropriate. This is certainly one argument made by judges to claim they are not subject to extra-legal influences the claims of political scientists notwithstanding (Kozinski [1993] 2013).

At the same time, there is reason to suspect heuristics persist at all stages of judicial conduct. Federal district court judges in particular are likely prone to adopt cognitive shortcuts in light of the challenges they confront. Initially, Federal district courts are no longer primarily trial courts, but rather the facilitators of resolutions short of jury trial. (Hornby [2007]2013). Limiting the number of trials through earlier engagement and active case management has been stressed for decades, leading to greater numbers of civil cases settling and the vast majority of criminal cases resulting in plea bargains. (Boyd 2013). The “disappearing jury trial” has been commented on less as a loss for democracy and Constitutional values than the seemingly inevitable shedding of a nuisance. (Hornby [2007]2013: 107-111]).

One area where district court judges have been forced to engage more actively with legal process at an earlier stage is reviewing and ruling on the admissibility of expert witness testimony. It is an area with a fairly well developed legal doctrine, but with sufficient ambiguity and highly deferential review on appeal, permitting judges significant leeway in how they approach the task. It is a required task, but one that frequently requires a judge to assess unfamiliar topics based on limited information and time, creating a measure of uncertainty as to what the correct answer may be. These conditions, incentivize minimizing the mental burdens involved in deciding whether to find the testimony sufficiently relevant and reliable¹ to admit it. (Simon 1955; Kahneman 2011).

To appreciate the significance of this task, it is worth exploring what an expert is and how judges came to have a responsibility to assess all expert testimony for reliability. Per Federal Rule of Evidence (FRE) 702, an expert is anyone possessing training, knowledge or experience that would “assist the trier of fact to understand the evidence or to determine a fact in issue.” As the rule suggests, there is no hard and fast dividing line between lay and expert witness, and no boundaries on the scope of potential expertise. They can range from a mechanic who can explain how a faulty repair made a car unsafe to a Ph.D. biochemist explaining how the

¹ In the law, as noted, reliability refers to the trustworthiness of a piece of evidence – can it be used as the basis for establishing an element of a case, under a set of norms developed over time that rely primarily on human experience and some insights from early research in psychiatry. It is not, as used in science, related to an instrument’s ability to generate consistent results. Moreover, at the trial court level the initial admissibility determination is not the same as a validity assessment in science – the ultimate judgment of the truth of the statement is left to the finder of fact, which may be the judge or a jury. Throughout this paper, the term is used in its legal sense, but the distinction points up in one small way the different world view of judges and social scientists.

reagents in a blood glucose monitoring strip work. The common thread is that their expertise will provide a different context or meaning for jurors than the raw information would.

Where a witness is speaking to something that is within their field of expertise and will assist the trier of fact, expert witnesses are free of some of the key constraints that limit the admission of other evidence. (FRE 702 *et seq.*; cf. FRE 602, 701). They are not limited to matters within their personal knowledge but can rely on the words and work of others. They can offer opinions, rather than just factual statements, so long as the opinions rest on their expertise. They can even offer opinions as to the ultimate question in a case, although they generally cannot testify in the form of a statement that explicitly tells the jury how to find on an issue like negligence. In many cases, expert witnesses are the linchpin of a party's presentation on everything from the fact of harm to the causation of harm to the amount of damages that resulted, a way of dramatically weaving together disparate strands of evidence into a complete picture. (Vidmer 1995).

The act of admitting an expert can itself be significant to case outcomes, as it appears to make their testimony more credible in the eyes of jurors than if the testimony is presented without such an endorsement. (Schweizer and Saks 2009). This suggests the jury is taking a cue from the judge that the material must be significant and to some degree true. Because of these factors, there has long been concern that jurors may be swayed by convincing but misleading expert testimony. (Schauer and Spellman 2013). There is an experimental basis for this fear – people generally assess experts, like other witnesses, based on issues like appearance, demeanor and background as much as the actual content of their testimony. (Vidmer 1995). Jurors often fail to identify or understand the value of validity problems with scientific evidence. (McAuliff,

Kovera and Nunez 2009; McAuliff and Duckworth 2010; Levitan 2017). Moreover, trying to provide scientific education in the guise of expert testimony is difficult - using additional witnesses to point out problems in testimony actually tends to result in test subjects distrusting both sides' experts rather than causing them to accept valid and reject invalid testimony. (Levett and Kovera 2007). Thus, there does seem to be a need for some measure of screening to minimize the potential for expert testimony to mislead.

By 1993, a conflict had arisen over how judges were supposed to assess expert witnesses. Many Federal courts nominally followed the “general acceptance” standard - if the expert’s testimony conformed to commonly accepted practices in their field it was typically admitted. (Beecher-Monas 2011). This arguably conflicted with the text of the Federal Rules of Evidence, as the judge was not independently assessing the reliability of the evidence. In *Daubert v. Merrill Dow Pharmaceuticals* (1993), the Supreme Court rejected this standard as too cramped in the context of scientific evidence. (506 U.S 579, 589). Instead of simply looking to issues of acceptance, courts were supposed to employ a multi-factored approach to assess the scientific rigor of the witness’ method derived from a Popperian conception of science, including falsifiability, error rate/replicability, the use of standards and the subjection of results to peer review as well as acceptance. (*Daubert* , 506 U.S. at 593-594; Jasanoff 1995).

Subsequently, in *General Electric Co. v. Joiner* (1997) the Court found that a trial court’s consideration of the admissibility of expert testimony should be reviewed under an abuse of discretion standard. (522 U.S. 136, 142-143). This essentially conceded that such judgments were too particularistic for reviewing courts to assess accurately except in extreme cases. Adding to the complexity of the situation, in *Kumho Tire Co. v. Carmichael* (1999) the Court

clarified that all proposed expert testimony should be analyzed for intellectual rigor, although the relevant factors that could vary depending on the nature of the testimony. (526 U.S. 137 at 146, 150). Together, these cases are referred to as the “*Daubert trilogy*”, the substantive standards are sometimes referred to as *Daubert/Kumho* standards, and collectively reflect the latest substantive word from the Supreme Court as to how trial courts are to assess the admissibility of proffered expert witness testimony.

To date, analysis of admissibility rulings has taken one of several forms. There have been a significant number of doctrinal analyses, which use unsystematically selected cases to assert trends in practice. There have also been several assessments of outcome looking at the impact or lack thereof of shifting from *Frye* to *Daubert/Kumho*. There have also been descriptive assessments of the conduct of judges, including some textual analysis of rulings and survey responses. But there has been only a limited effort at assessing the actual decision-making process, and that has relied on outcomes rather than the actual text of the rulings in keeping with the bulk of empirical social science on judicial decision making. Thus, there remains multiple levels of mystery about the admission of experts – we have some indicators of what judges are doing, but far less of an explanation of why.

There is some evidence suggesting that a heuristic model explains the process as well or better than other explanations. Gatowski, Dobbin, Richardson, Ginsburg, Merlino, and Dahir (2001) report that judges understand certain *Daubert* factors – general acceptance and peer review – much better than they do more complex issues such as replicability and error rate. Dixon and Gill (2002) find that general acceptance is the *Daubert* factor judges are most likely to analyze, and that it was predictive of judicial conduct in most cases. This is also consistent with

Merlino, Springer and Sigillo's (2011) finding that general acceptance is a key factor in the determination of admissibility, and with Helland and Klick's (2012) finding that courts expressly using general acceptance as the controlling standard have similar numbers of testifying experts as courts applying the flexible multi-factored standard of *Daubert*. This suggests that judges likely consider general acceptance or something akin to it as a threshold or benchmark consideration before assessing whether to investigate other, more complex *Daubert* factors. This would also be consistent with jury studies showing jurors tend to fall back on "peripheral processing" issues like signals of trustworthiness or sincerity, rather than the scientific content of expert testimony, particularly in complex cases. (Cooper and Neuhaus 2000). For judges, general acceptance may well replace cues like body language, dress or perceptions of the speech or attitude of a witness relied on by jurors. However, expert witness admissibility has not been explored with a distinct focus on whether judges are adopting a heuristic approach in making these determinations.

Another heuristic for judges related to admitting experts may lie in deciding whether they need to intervene at all. Cecil and Willigang (1993) report that judges are reluctant to be seen as intervening in the adversarial process through measures such as the appointment of expert witnesses. Both *Daubert* and the Advisory Committee notes to Rule 702 stress the availability of actions including cross-examination, presentation of countervailing witnesses and jury instructions as means to address weak expert witness evidence without excluding it. The Advisory Committee also asserts admission should be the rule, rather than the exception. This may well bias judges in favor of avoiding taking a hard look at complex questions of methodology, so long as they believe the procedural fairness of the adversarial process is not

threatened. Thus, there is a strong possibility that analyzing judicial opinions regarding the admissibility of expert witnesses both qualitatively and quantitatively will support a finding that judges are using multiple approaches that minimize the cognitive effort required to assess these potentially complex questions.

II. Data And Methods

In keeping with this evidence, this paper presents a portion of the results of a non-experimental, cross-sectional comparative research study analyzing Federal trial court determinations of the admissibility of expert witnesses in civil cases. The study relies on thematic analysis of judicial rulings on these issues from nine United States District Courts over a six year period (rulings issued from 2010 through 2015).

This study sought to overcome a number of methodological gaps and questions raised by prior studies regarding judicial decision making generally and the treatment of expert witnesses in particular. Initially, by looking not just to outcomes but the content of judicial decisions, the study sought to get a fuller understanding of the mental processes of judges than could be achieved by solely investigating the outcome as it related to judicial characteristics such as gender or politics. At the same time, it avoided the risks of faulty or biased recall associated with interviews and should serve as a conservative indicator of heuristic tendencies. Additionally, by capturing all of the rulings within a constrained set of districts, the study sought to overcome the potential for distortion from assuming uniformity across different geographic jurisdictions. Employing qualitative thematic analysis permitted the author to identify heuristic tendencies in the data that had been overlooked by prior research, such as the adoption of

simplified standards for assessing challenges and how judges tend to restrict the testimony of experts rather than exclude them completely even when challenges are upheld.

One of the goals of the overall study that this paper stems from was to ensure that it captured the potential distinctions at the circuit level in a meaningful manner. Preserving this potential diversity of practice while avoiding supported a purposive sampling approach to case selection. (Seawright and Gerring 2008: 300-301). At the same time, there was a concern that inclusion of certain atypical districts – such as the Northern District of Illinois, which is dominated by Chicago and dwarfs other districts in the Seventh Circuit, or the Eastern District of Texas, which has become a center for patent filings to the exclusion of much of the rest of the country – would fundamentally skew any data. Instead, nine district courts - 3 geographically adjacent districts in each of the Ninth, Seventh and Fourth United States Courts of Appeals – were selected. In each case, the cluster included a mixture of urban and rural areas, so as to preserve a broad cross section of potential litigation. They included the District of Oregon, Western District of Washington and Eastern district of Washington within the Ninth Circuit, the Northern District of Indiana, Southern District of Indiana and Southern District of Illinois in the Seventh Circuit, and the Eastern, Middle and Western District of North Carolina in the Fourth Circuit.

While preserving the potential for unique potential appellate influences, the research was designed to try and avoid geographic or case type dissimilarities. As noted above, courts with a unique, dominant type of case were consciously excluded. Each cluster also provided a similar geographic cross section. Each contains one district dominated by major metropolitan region – Seattle in the Western District of Washington, Indianapolis in the Southern District of Indiana

and Charlotte in the Middle District of North Carolina. Each contains a district that includes one significant but secondary city with a large population and is otherwise more scarcely populated - Portland in the District of Oregon, Fort Wayne in the Northern District of Indiana and Raleigh in the Eastern District of North Carolina. Finally the remaining district in each cluster (the Eastern District of Washington, Southern District of Illinois and Western District of North Carolina) lacks any large cities although each has some regional population centers such as Spokane, Washington, Springfield, Illinois and Asheville, North Carolina.

Given the focus of this study on the typical conditions under which rulings are made, the goal was to identify every written opinion issued from 2010-2015 that involved the admission of expert testimony in a civil case within each district under study. To attempt to ensure all possible opinions were included, initially, the PACER (Public Access to Court Electronic Records) system was employed to identify opinions issued from the District of Oregon during 2015 on a month by month basis. This system lacks a key word search or similar function, so the cases embracing a *Daubert* issue were identified by manual review. The cases so identified were then compared with the results of a LEXIS search for the District of Oregon for the same time period that specifically mentioned “*Daubert*” or “Rule 702.” There was complete capture of the former set in the latter search, creating confidence that LEXIS could be used to compile the datasets for this project.

Three datasets were compiled for initial qualitative analysis. A set of LEXIS searches designed to maximize the capture of potentially relevant cases were conducted. In each case the search was for the same date range of 2010 to 2015 and limited to cases referencing the districts of interest. The initial search was for all cases mentioning the terms *Daubert*, *Kumho* or the

number 702 within 25 words of the term expert. This was in turn narrowed to cases with one of the terms object, exclude, strike or limit within 10 words of the term expert. Cases were then reviewed to exclude all non-district court rulings (including both appellate court and bankruptcy court proceedings), criminal cases, those that were from irrelevant districts, were duplicative or lacked any substantive analysis of the question of expert admissibility (such as scheduling orders, instances of summary adoption of a magistrate's recommendation or cases in which the ruling was held in abeyance pending trial). This resulted in 292 identified cases with substantive discussions of challenges to the admission of expert testimony.

The cases were originally coded individually within atlas.ti. The focus was on the content, context and outcome of the cases at this stage. Thematic analysis was employed to review and assess the material. (Braun and Clarke 2006). Beginning from the framework of directed content analysis, initial codes were derived from prior theory and empirical work as a basis for comparison. (Hsieh and Shannon 2005). At the same time, items of interest that developed during review were also coded to provide a more robust understanding of the data. Both during and upon completion of the coding, the results were analyzed for the identification of overall themes that emerged from the process. These themes are summarized in table 1 and discussed in the following sections of the paper.

Theme 1	Excluding an Expert is an Unusual Step/Admission is the Default
Theme 2	Judges Rely on Socially Constructed Standards to Assess Methods and Experts
Theme 3	Judges' Rulings Demonstrate Significant Cognitive Simplification

Table 1: Themes Identified in Qualitative Analysis

III. Results and Discussion

Thematic analysis is a fitting technique for analyzing district court opinions in light of the combination of isomorphism and individuality expressed in them. Judges tend to follow a similar style in ruling on motions such as the requests for exclusion studied here. They start with a recitation of the basis of the lawsuit, a summary of the motion or motions that are to be resolved, a brief statement of relevant facts and legal standards, and an application of the legal standards to the fact. As discussed in greater detail below, this is the first of several indications of cognitive simplification by judges – the creation of what amounts to a checklist or scaffolding for the legal analysis allowing reflexive rather than reflective assessment (Kahneman 2011: 20-22). In itself this isomorphic tendency among judges is not necessarily problematic, but it is suggestive of both potential problems and ultimately some approaches to minimizing the risks of this practice.

This is not to say that opinions are entirely formulaic or monolithic. Some are incredibly detailed in description, some almost unbelievably brief. For an example of the former, Judge David Herndon of the Southern District of Illinois constructed a detailed, multi-paged overview of expert witness law to apply to a number of rulings in the multi-district litigation relating to the marketing of Yasmin/Yaz brand birth control. *See In re Yasmin and Yaz (Drospirenone) Marketing, Sales Practices and Products Liability Litigation* (S.D. Ill. 2011), 2011 U.S. Dist. LEXIS 145552. He then tailored it depending on whether treating physicians, social scientists or other particular subtypes of expert were involved. *See id* at * 12 (adding discussion of treating physicians); cf. *In re Yasmin and Yaz (Drospirenone) Marketing, Sales Practices and Products Liability Litigation* (S.D. Ill. 2011), 2011 U.S. Dist. LEXIS 145593, *13 (adding discussion

specific to social scientists); *In re Yasmin and Yaz (Drospirenone) Marketing, Sales Practices and Products Liability Litigation* (S.D. Ill. 2011), 2011 U.S. Dist. LEXIS 145989 (lacking either). By comparison, Judge Marsha Pechman of the Western District of Washington managed to resolve an expert challenge in an order totaling four paragraphs – yet still retaining an overview of the parties, the nature of the challenge, the law, and its application. *Oswalt v. Resolute Industries* (W.D. Wash 2012), 2012 U.S. Dist. LEXIS 190123.

Other distinctions exist, too, within this general paradigm. Some announce the outcome of the motion in the very title of the document – “Order Granting Motion to Exclude” for example – whereas others place it in the opening paragraph or hold the result for the very end of the document. *Dasho v. City of Federal Way* (W.D. Wash. 2015), 101 F.3d 1025, 1027; *cf. Fuentes v. Miller* (N.D. Ind. 2015), 2015 U.S. Dist. LEXIS 82978, *Agrigenetics v. Pioneer Hi-Bred International, Inc* (S.D. Ind. 2010), 2010 U.S. Dist. LEXIS 12003. Most attempt to project a tone of seriousness, but sometimes the personality of a judge will shine through in wordplay. Perhaps the clearest case of this in this study came in a case involving a dispute over agriculture chemicals of all things. Magistrate Tim Baker described the each of the parties in *Agrigenetics* as “turning its parasite–eliminating expertise toward eradicating the other side’s experts” and in denying the motions to exclude finding “[t]he parties’ experts have proven more resistant to attack than seed corn parasite.” 2010 U.S. Dist LEXIS 12003, * 2, 16. But these differences are secondary to the overall points of commonality.

Looking at these common points as a means of understanding the decision making process revealed several dominant themes that suggest judges are relying on heuristics in resolving these issues. The first was that, consistent with the Advisory Committee notes to

Federal Rule of Evidence 702, exclusion of an expert is the exception, while admission is the rule, across all types of expert and cases. The choice architecture of the controlling law has created what may be called the “admissibility heuristic.” The second is that judges do not appear particularly comfortable assessing scientific and technical methods directly, and would rather rely on the judgment of third parties or the collective judgment of prior cases where possible. This tendency to substitute the judgements of others for independent analysis can be called the “acceptance heuristic.” How courts tend to sort objections and process cases suggests a third and final theme – that judges have tremendously simplified the *Daubert/Kumho* framework for assessing complex challenges across fields, and in so doing reconciled the process with other types of evidentiary rulings. In this paper that tendency is referred to as the “framework heuristic.”

All of these simplifications are logical steps, and in most cases are likely no worse than the alternative. However, as with the general field of heuristics and biases described by Kahneman (2011), there is reason to believe that these lead to suboptimal outcomes on some occasions. This in turn suggests policymakers should consider limiting the potential adoption of heuristics when proposing changes to judicial conduct and ways that practitioners can seek to use the heuristic tendency of judges to their advantage.

A. The Admissibility Heuristic - Excluding an Expert is an Unusual Step

The literature on the impact of *Daubert* is at times conflicted. On the one hand, reading Buchman (2007: 681-85) one would get the impression that the default standard for a conservative judge is to exclude experts when challenged, whereas liberals were slightly more likely to admit an expert than not. Merlino, et al. (2011: 9-15) found no default to exclusion

generally, but their finding about the overwhelming rejection of toxicology experts in their sample suggests there should at least be some highly disfavored types of testimony. At the same time, neither of those reports comport with Helland and Klick's (2012: 17-31) finding that cases in both state and Federal courts have similar numbers of experts. Nor is it in keeping with Cecil and Willigang (1993: 5, 20-21), who found most judges wanted to respect the adversarial process rather than intervening with regards to expert witnesses. In fact, both the text of the *Daubert* and *Kumho* opinions as well as the Advisory Committee notes to the revision of Rule 702 in 2000 that codified *Daubert/Kumho* noted that exclusion is the exception, not the rule. Instead, both the cases and the rule point out the availability of actions including cross-examination, presentation of countervailing witnesses and jury instructions as means to address weak expert witness evidence without excluding it.

Ultimately, it appears that in the great majority of cases judges are at least attempting to honor the statements in the Advisory Committee notes to Rule 702 to use exclusion sparingly. Although the primary focus of this paper is the content of rulings, even a cursory look at the rate of exclusion makes this distinction clear. Where Buchman found the mean admission rate in all cases was just under 44 percent (Buchman 2007: 681), the overall admission rate in this sample was in excess of 82 percent. When controlling for region, there was some variability, with the Ninth Circuit issuing a rejection in 25 percent of cases where the Fourth and Seventh Circuit were issued a rejection in less than 16 percent of cases, but nowhere approached the frequency of rejection reported by Buchman (2007). Judges simply default to admission, in keeping with *Daubert* and Rule 702.

As a general rule, judges appear to want to admit an expert when they feel there is a plausible justification of the expert's methodology. As will be discussed in Part B below, this often leads to a "safe harbor" effect when experts suggest they are comporting with an industry norm that can short circuit the analysis. At the same time, even when the expert deviates from the industry practice a good explanation will likely lead to admission rather than rejection.

3600 Michigan Co., Ltd. v. Infra-Metals, Co (N.D. Ill. 2011) 2011 U.S. Dist. LEXIS 574 is a good illustration of this judicial tendency. The parties disputed whether a landlord had made a good faith effort to secure a substitute tenant and thus mitigate the harm of termination. To try and buttress its assertions, the plaintiff hired a real estate appraiser to assess the market in the relevant time period and establish if the asking price was reasonable. The expert admitted he did not use the leading industry standard for appraising property in reaching his conclusions – which the defense seized on as the basis for attacking his testimony.

This might seem like the paradigmatic "easy case" – and admissions of variance or lack of rigor are often part of a successful challenge. However, Magistrate Judge Andrew Rodovich was willing to give the expert the benefit of the doubt when there was a plausible justification for his conduct. Here, it had to do with the disconnect between the client's need in the case – for a general survey of market conditions – rather than a focused market based appraisal of a particular property. Given that the leading standard had no specific guidance on such a task, and that the expert used techniques such as interviews and personal knowledge typical of similar studies, the court ultimately decided that the deviation was immaterial to the reliability of his methods.

Even where the judge finds that the core of an expert's testimony is improper, they are often loath to exclude the witness entirely. *Dasho* illustrates just how far a restriction, rather than an outright rejection, can go in limiting a party's case. This case involved assertions of excessive force by police officers in Federal Way Washington. Specifically, the plaintiff was shot several times by multiple officers while wielding what Judge James Robart somewhat cryptically referenced as a "blunt tipped knife" while naked and highly agitated. 101 F. Supp. 3d at 1027.

Although he was convicted of criminal assault for his conduct with the knife, Mr. Dasho sued on the grounds specifically that the shots were unjustified and more generally that the entire encounter was excessive. He retained a forensics expert who asserted he could reconstruct the sequence of events, including the precise sequence of shots fired, in a way that contradicted the officers' sworn statements. Mr. Dasho also retained an expert in "police procedures" who would argue that the conduct in question was excessive. Both were challenged by the defendants, although the primary focus of their claims was the shooting narrative expert, Kay Sweeney.

Judge Robart' opinion makes it clear that either Mr. Sweeney or Mr. Dasho's lawyers did not understand their obligations in writing an expert report. The judge had already found the initial report highly problematic:

"In its prior order, the court observed that Mr. Sweeney's report consists mainly of a list of the physical evidence followed by Mr. Sweeney's conclusions about shot sequence and Mr. Dasho's movements. . . . The report contains little reasoning connecting the Opinions to the physical evidence and offers no information to support the reliability of the Opinions. . . . Based on that inadequate record, the court could not determine whether Mr. Sweeney had employed reliable methods and applied those methods reliably to the facts."

101 F. Supp. 3d at 1031 (internal citations omitted).

Despite all these problems, Judge Robart had not excluded the witness as a discovery sanction, but requested a supplemental filing “detailing Mr. Sweeney’s methods, their reliability and how Mr. Sweeney applied those methods to this case.” *Id.* What the judge got, though, was none of that – instead, there were a number of conclusory assertions by counsel the methods were “universally accepted” without citation to any source. There was no direct explanation of the method, but instead a trial transcript that suggested the method for establishing the shot pattern was Mr. Sweeney’s subjective creation from physical evidence of a “sequence of events that seems to him to be the best overall match.” 101 F. Supp. 3d at 1032.

Not only could Mr. Dasho’s counsel not support their claims with a single, citation, Judge Robart stated that despite doing his own research he could not find “a single instance of a federal court allowing a forensic expert to offer opinions of this type based on a like methodology.” 101 F. Supp. 3d at 1032. Instead he found several cases rejecting the idea of reconstructing a shot sequence based on such items of evidence as shot trajectories, entry and exit wounds. If that was not bad enough, it turns out Judge Robart was not writing on clean slate with regard to Mr. Sweeney. His methods had been assessed more than a decade earlier by the Washington Court of Appeals and found to be wanting in any sort of rigor. 101 F. Supp. 3d at 1033. For these reasons Judge Robart found the methods unreliable. Moreover the judge found the methods such as they were had not been applied in a way that could be assessed for reliability – there was just no way to even assess many of the claims that were made.

The plaintiff’s police practices expert, a former police Chief named D.P. Van Blaricom did not fare much better. He was to testify specifically as to how the use of force in the particular case compared to police standards. However, either the expert or counsel were not

comfortable actually asserting that the conduct in question was more probably than not excessive force, but only that it “may have been.” 101 F. Supp. 3d at 1035. Unfortunately for the plaintiff, “may have been” standing alone is irrelevant, as it falls to meet the standard of liability. As a result, the court excluded this element of the witness’ testimony.

What he did not do, though, was exclude either witness in their entirety. Even though Judge Robart had seemingly gutted the most important elements of the plaintiff’s case, he didn’t find that the two witnesses could not testify in their entirety. While Mr. Sweeney’s shot sequence reconstruction evidence was excluded he had also provided a number of foundational statements about shot trajectories, wound patterns and locations that would be permitted to come in. Likewise Mr. Van Blaricom was not excluded in his entirety – he’d still be able to testify to general police procedures that would help the plaintiff establish the scene for their version of events and how the police departed from established norms.

This reluctance to exclude experts is understandable at several levels. It is not only consistent with the Advisory Committee notes and statements in *Daubert/Kumho* about defaulting to admission, but also with general legal norms that extend beyond the four corners of the *Daubert/Kumho* line of cases. Although judges are cautioned in *Daubert* to worry about the potential for experts misleading a jury, they are also fundamentally committed to the value of the adversarial process. That not only includes a belief in the ability of lawyers to correct the record through cross-examination and alternative witness, but also a view that their own role is to be limited to policing the parties’ conduct rather than intervening (Cecil and Willging 1993: 18-21).

The cases reviewed in this study suggest that this perspective underlies many of the rulings judges arrive at, and tends to make them skeptical of the need for intervention. On 135

occasions judges remarked on the ability of cross examination to be used to address alleged deficiencies in the methods or conclusions. This was one of the two most frequently cited bases for rejecting a challenge. The only one that exceeded it, and was frequently cited along with it, was the concept that a particular challenge went to the weight to be given the evidence, rather than the admissibility of the expert.

Moreover, the cases in the dataset suggest why judges think the adversarial process works so well. When experts were rejected entirely it was often attributable to the adversarial process – either their own testimony or conflicts between what that testimony and various assertions of counsel drawn out through careful cross-examination. Thus, John Lamberth, a social psychologist who had discussed the difficulty of accurately assessing the racial disparities in law enforcement from passive observation limited to traffic violations nonetheless adopted these techniques in a racial discrimination case and claimed they were valid – a claim Judge Thomas Schroeder found lacking in credibility. *United States v. Johnson*, 122 F. Supp. 3d 272 (M.D. N.C. 2015). It did not help that the methods of single observer, fixed observations he had relied on were actually criticized in the very articles he pointed the court to as supporting his techniques, and that he essentially relied on subjective understandings of visual indications of Hispanic ethnicity with no apparent effort at meaningfully testing those approaches. All of this combined to lead to the exclusion of his testimony, which in turn made it virtually impossible for the Federal government to maintain its assertions of racial discrimination.

A similar fatal admission was made by Stephan Neese, a purported accident reconstruction expert in the matter of *Stachon v. Woodward*, (N.D. Ind. 2015), 2015 U.S. Dist. LEXIS 129958. The key question in the case was the location of a pedestrian when they were

struck by a truck. Neese argued that the impact was between 100 and 130 feet from the final resting place of the plaintiff based on the location of a single sock. Yet he admitted under cross examination that clothing was an inappropriate marker and had no explanation for why he thought one sock was reliable but another located more than 60 feet away was not. This, when coupled with the fact his “reconstruction” used this assumed distance, an unrepresentative truck and a wooden model that scaled out to a 6’9” human and an impact distance 60 times greater than observed led Magistrate Judge Andrew Rodovich to reject the testimony in its entirety.

Extreme examples like this and similar situations likely cause judges to overestimate the value of cross-examination and contrary testimony. Judges are trained from law school and practice to think of cross-examination as the key skill at impeaching the testimony of witnesses, limiting or destroying their value. Similarly, counter witnesses are stressed as giving the jury a choice and allowing them to accept the side they find most credible. When judges are moved to rule it is often because of these tools and how they highlighted particular gaps in testimony, both in the context of expert witness and elsewhere.

This in turn may trigger the cognitive shortcut known as the availability heuristic. As Tversky and Kahneman (1974: 1127-1128) pointed out, generally speaking frequent events are easier to recall, and thus what is most readily recalled can often be relied on as an initial estimate of likelihood. However, when unusually salient or exceptional events are judged against those that are more mundane, the exception may be thought of as more likely to occur than the mundane event because of the ease of recall. For judges, their entire careers involve looking for and ruling on questions contested through the adversarial process. Instances of highly effective cross-examination or arguments where the judge had confidence in the correct nature of their

ruling likely abound, and are apt to be highly salient in light of a judge's training and experience. Thus, it is easy for a judge to think that these same tools will be highly effective if and when a jury is exposed to them. This may explain the prevalence with which judges note that an objection goes to the weight to be given evidence by the jury rather than its admissibility and the ability for the matter to be explored through cross-examination.

While these sentiments are thus understandable, empirical studies suggest cross-examination and contrary witnesses on technical matters are not very efficacious in helping jurors arrive at an accurate understanding of issues (McAuliff, et al. 2009: 252-253; McAuliff and Duckworth 2010: 495-497). Initially, witnesses often cannot spot significant errors that would raise validity concerns about testimony, with and without cross-examination. When subject matter is so complex as to be beyond juror's knowledge to process, their reaction is often to focus on peripheral issues such as how much the witness is paid or their connection to the district as a means of assessing their general trustworthiness (Cooper and Neuhaus 2000: 168-171). This limits the possibility that cross-examination on technical questions is apt to be effective at either educating jurors or swaying them. Moreover, the introduction of a rebuttal witness who noted methodological questions was not to lead jurors to evaluate the study more carefully, but to be skeptical of all expert testimony in a "pox on both your houses" effect (Levett and Kovera 2008: 369-370). The latter impact might benefit defendants, as the neutralization of a plaintiff's expert would typically undermine their ability to support critical elements it must establish to prevail. However, it certainly does not further the aims of the adversarial system.

Thus, while the tendency to admit experts is consistent with legal norms it is also consistent with heuristic processes. An alternative construction of *Daubert/Kumho* and the

Advisory Committee notes about liberalizing admissibility, along with the standard of review being set at an abuse of discretion in *General Electric v. Joiner*, is that it creates a “choice architecture” in the words of Thaler and Sunstein (2008: 6) that “nudges” judges towards admissibility. As Baum (2006: 50-53, 99-100) notes and Kozinski ([1993]2013: 117) echoes, professional reputations matter to judges. Being reversed can be a blow to that reputation. It also typically means that the same judge will have to repeat prior efforts, which is at odds with the pressures courts are under to settle and resolve cases promptly. It is a far less risky option to admit an expert from the perspective of a judge concerned with avoiding reversal. This may well bias judges in favor of admission in normal cases consistent with the precepts of prospect theory as set forth by Kahneman and Tversky (1979: 277-280, 289-90). When coupled to the underlying concepts of the adversarial process, the judge is given a series of simple shortcuts to avoid wading into complex questions in the vast majority of cases – admit and admonish the parties to use the adversarial process to address most issues.

Ultimately, it appears that judges are at least attempting to follow the edicts of the law to simultaneously demand intellectual rigor from experts while applying the standard for admission liberally. In doing so, however, they are also following the expectations of cognitive theories of decision making. Admissibility is a heuristic in this area of law, with judges inclined to find a way to admit the testimony if at all possible. As with most heuristics, however, the simplification of the judge’s task can lead to sub-optimal outcomes, specifically in terms of fully analyzing potentially problematic experts.

B. The Acceptance Heuristic - Judges Substitute Third Party Standards and Prior Cases for Their own Assessment of Methods and Particular Experts

In most models of judicial decision making, there is an implicit assumption that the judge's conduct is predicated on a thorough analysis of the case's substance. This assumption does not mesh well with information on how people generally assess complex questions. As Simon (1955: 101-105) pointed out in one of the earliest challenges to rational actor theory, there are very few circumstances in which an individual will truly engage in comprehensive assessment of all the information that is potentially relevant to a decision. More often they will establish parameters and adopt a solution that "satisfices" those parameters. So long as the minimum threshold is met, in most cases decision makers stay within their boundary conditions, as the cost of further information or assessment is not justified.

Prior empirical assessments of rulings on expert witnesses suggest that there is some degree of this type of satisficing being practiced by judges. Dixon and Gill (2002285-288, 299-300) noted that judges seemed to emphasize whether an expert's testimony was of a type "generally accepted" in the field, despite *Daubert*'s indication that this standard from *Frye* was not controlling. Merlino, et al. (2011: 17) similarly found that general acceptance was regularly cited. This is also consistent with Helland and Klick's (2012:32-33) finding that admission rates appeared to be similar whether a court only applied the simple and allegedly narrower *Frye* standard or followed *Daubert*.

This tendency is understandable. Judges evaluating expert witnesses are routinely asked to evaluate a broad range of esoteric topics. In one case, they may looking at something as seemingly trivial and yet highly technical as the proper design of a sidewalk and its height relative to a street. *Delarosa v. Speedway, LLC* (S.D. Ill. 2013), 2013 U.S. Dist. LEXIS 125104. The next case in the pipeline may, like *Dasho*, require an evaluation of novel approaches to

shooting reconstruction. A third dispute, such as the *In re Yasmin and Yaz* MDL cases, might span everything from specific and general causation to labeling requirements to medical monitoring efforts – there were 16 experts submitted on liability related issues alone. A fourth might raise questions about how to evaluate supposed expertise on the “culture of NASCAR” in a contract dispute over sponsorship. *See Team Gordon Inc. v. Fruit of the Loom, Inc* (W.D. N.C. 2010) 2010 U.S. Dist. LEXIS 14227. Yet another might call on the court to grapple with questions of how to handle experts on intersectionality, group pressure and implicit bias among members of traditionally disadvantaged groups. *See Apilad v. North American Gay Amateur Athlete Alliance*, (W.D. Wash. 2011), 2011 U.S. Dist. LEXIS 159575. No one could realistically be expected to understand the nuances of such a diverse array of fields of expertise in the compressed time afforded district courts to resolve these disputes.

Instead, judges frequently look to third party standards as a means of assessing the validity of the particular method being challenged. These are frequently drawn from industry manuals or from prior legal evaluations of similar methods. Deviation from a standard is not fatal if there is a good explanation for it. Moreover, in the case of a developing technology a court might be more apt to overlook a lack of consensus or clear standard in the field. *See Ruppel v. Kucanin*, (N.D. Ind. 2011), 2011 U.S. Dist LEXIS 167505 (discussing and permitting testimony as to results of diffusion tensor imaging or DTI for diagnosing brain injury). However, where they do exist they create a “safe harbor” or “legal science” effect that can potentially distort the value of evidence and make it either unduly difficult or extremely easy to admit (Jasanoff 1995:130-131; Jasanoff 2008: 128-129; Beyea and Berger 2001: 348-360; Beecher-Monas 2011). Although there has been a movement by the Federal Judicial Center

(FJC) and the National Academies to align scientific understanding with legal requirements through a reference manual on science generally as well as specific fields of science, its relatively infrequent citation suggests that it has yet to achieve the same impact as the social construction of industry bodies or court cases.

In this study, two fields that came up with some frequency and highlight the issues of these safe harbors were fire investigation and medical causation testimony couched in terms of “differential diagnosis” or “differential etiology.” As to the former, the frequent citation of the standards of the National Fire Protection Agency’s standards reveals that it does not appear to take much substance to create a widely held standard. As to the latter, their frequent use even as the FJC Reference Manual on Scientific Evidence has sought to clarify that these terms as used in law do not reflect medical standards raises questions whether courts are truly honoring *Daubert*’s gatekeeping process in many cases. This problem is further underscored with the recognition in the medical community, as reflected in the latest edition of the FJC Reference Manual on Scientific Evidence that statements of certainty by medical professionals should be taken with caution given the different values underlying medical care and legal testimony.

The National Fire Protection Agency is a trade group that, among other things, writes fire codes that can in turn be adopted by local governments. In 1992, it sought to create a standard for fire investigation that created a formal method (Watson 2015: 24). It was modeled on basic principles of the scientific method including observation, evidence assessment, alternative consideration, hypothesis development and testing. It was reportedly quite contentious among practitioners at the time of its adoption because of its use of formal language and the fact that it was presented as a consensus approach to fire origin investigation (Watson 2015: 25). Over the

ensuing two decades regular revisions have occurred and despite complaints at times from field practitioners of investigation it has become a widely accepted industry standard.

This acceptance has in turn led to many court cases that treat the invocation of NFPA 921 as creating a default presumption of admissibility for fire investigations. Emblematic of this is *Citizens Ins. Co. of the Midwest v. LG Electronics USA, Inc* (S.D. Ind. 2012), 2012 U.S. Dist. LEXIS 1127128. There Judge Richard Young denied a motion to exclude expert testimony in a case where fire origin and causation was at issue largely predicated on finding that the fire origin expert had adhered to the tenets of NFPA 921. He began his discussion by finding that NFPA 921's general outline of investigation processes was a reliable method. He then proceeded to compare that general method to the assertions of the experts in question and found that they had generally followed the method.

What is remarkable about this discussion is how general of a description of an investigation process was seen as sufficient to deem NFPA 921 a reliable method by Judge Young. Citing the 2008 edition of NFPA 921, he noted:

Chapter 4 outlines a basic scientific methodology known as the "systematic approach" to investigating fires (*Id.*, § 4.1). "With few exceptions, the proper methodology for a fire or explosion investigation is to first determine and establish the origin(s), then investigate the cause: circumstances, conditions, or agencies that brought the ignition source, fuel, and oxidant together." (*Id.*). Such investigation requires an examination of the scene, interviewing witnesses, and testing the results (*Id.*, § 4.3.2). The empirical data collected, which is "based on observation or experience and is capable of being verified," is subject to an analysis premised upon inductive reasoning (*Id.*, §§ 4.3.3, 4.3.5)

2012 U.S. Dist. LEXIS 112728, *7

Essentially, this is a basic overview of inductive reasoning and nothing more. It does not suggest why or how the investigator can be assured that their results are valid from following this

generic process. It offers no epistemological standards to ensure that an investigator's assessment is truly reliable. Yet, because “[c]ourts throughout the country have recognized that NFPA 921 offers a comprehensive, peer-reviewed, and detailed guide for fire investigation, and have held that its methodology is reliable for purpose of Rule 702” Judge Young followed suit without further examination. *Id.*

Similarly, in comparing the outline of the method to the challenged opinions, Judge Young seemed willing to indulge a very broad application of the method by the expert. Stephen Cottingham testified that he followed NFPA 921 by making observations, conducting interviews and constructing a hypothesis predicated on his process. However, his testing of the hypothesis was hardly what might be seen as rigorous in most fields – in particular, he “exchang[ed] ideas with other investigators on a more hypothetical level – i.e. ‘how can we explain these facts to either support or to no support a [] given theory?’” 2010 U.S. Dist. LEXIS 112728, *14 (brackets in original, internal citation omitted); *see also Severn Peanut Co., Inc. v. Industrial Fumigant Co.*, (E.D. N.C. 2014), 2014 U.S. Dist. LEXIS 34507, *6-7 (rejecting challenge to expert methodology based on NFPA 921’s endorsement of thought experiments as adequate process to assess hypotheses). Nonetheless, this was deemed to be conforming to an adequate degree by Judge Young, and he dismissed the failure of Cottingham to perform a rigorous “depth of char” analysis – something he acknowledged the NFPA itself labels ““the most reliable for evaluating fire spread”” and ““the key to generating reliable data”” – as an oversight that could be explored on cross-examination. 2010 U.S. Dist. LEXIS 112728, *13-14.

This is not to say NFPA 921 is not actually reliable. In its entirety it is over 400 pages long and includes a number of very specific guidelines such as the “depth of char” analysis.

However, its invocation appears to have become talismanic for many judges. This underscores how, in furthering the admonition in *Daubert* and the Advisory Committee Notes to Rule 702 to default in favor of admission, judges are potentially avoiding an actual assessment of the method. This at least creates the potential for the creation of a “legal science” subfield that gives unwarranted weight to techniques that are neither as certain nor as widely accepted as the testifying expert in an adversarial setting represents (Jasanoff 1995:130-131; Jasanoff 2008: 128-129; Beecher-Monas 2011: 60-62).

A second example of this phenomenon can be seen in the area of medicine, and in particular the process of establishing specific causation through the testimony of physicians. Generally, in a case predicated on exposure to a potential causal agent the plaintiff has to establish both general causation – that is, is a causal chain possible under any set of circumstance – and specific causation – that is, do the particular circumstances related to the plaintiff’s injury support a finding of causation. Physicians are often used to establish this second step. *See Koho v. Forest Labs, Inc* (2015 W.D. Wash). 2015 U.S. Dist. LEXIS 180860. In fact, under some circumstances – such as a very close proximity in time or a lack of plausible alternative causes – a physician’s testimony on likely specific causes might satisfy both general and specific causation. *See McClellan v. I-Flow Corp.*, (D. Ore. 2010) 710 F. Supp. 2d 1092 (finding that joint damage was so close in time and so unique relative to installation of a pain pump as to render distinct evidence of general causation unnecessary).

Koho illustrates why physicians have become a favored source of expert testimony on specific causation, particularly in the context of mass torts such as product liability cases. This case was one of a number of individual cases brought against the makers of Celexa for failing to

adequately warn patients and physicians about the potential risks of the drug, and in particular the risk of increased suicidal ideation. While not a class action, it was consolidated in the Eastern District of Missouri for seven years as part of MDL proceedings. In that time, one of the key common questions was general causation, which the MDL court found could be established through the expert testimony of David Healey. The same expert then submitted a tailored version of his expert report to Judge Lesnik on remand that “briefly discusses Mr. Ilich’s case” and “concluded that the drug ‘contributed significantly to [Ilich’s] death by suicide’ among other statements. 2015 U.S. Dist. LEXIS 180860, * 3 The basis for this finding was primarily accomplished by considering and eliminating another drug predicated on his view that there was not a “strong relationship with suicide” and the other drug, and the timing of the prescription of Celexa and the ultimate suicide. 2015 U.S. Dist. LEXIS 180860, *11. Judge Robert Lasnik ultimately found “Dr. Healy’s methodology sufficiently reliable for his testimony to be admitted at trial” and in turn predicated a denial of summary judgment for the defendant based on this finding. 2015 U.S. Dist. LEXIS 180860, *15; *Koho v. Forest Labs, Inc* (W.D. Wash 2015) U.S. Dist. LEXIS 46050.

Although Judge Lasnik noted that an utter failure to explain a conclusion would have supported the rejection of a claim predicated on differential diagnosis, the opinion makes clear that it is largely deferring to professional judgment. Essentially, because Dr. Healey is an expert in the impacts of SSRI drugs and their links to suicide, his statements that he did not see the same linkage with the other drug as he did with Celexa and that the timeline allowed him to rule out the deceased’s depression as the most likely cause were taken at face value. It also takes the validity of differential diagnosis as a general method largely for granted, predicated on prior

cases both within the Ninth Circuit and from outside in the particular context of suicides allegedly triggered by certain anti-depressants.

Judge Lesnik is far from alone in this - as with NFPA 921, differential diagnosis has become a shorthand for a presumably acceptable form of testimony. At the outer limits it is possible for a physician to be disqualified despite asserting to follow differential diagnosis, but typically only if the expert states they follow a standard that it can be shown they have deviated from without explanation. *See Messenger v. Norfolk Southern Railway Co* (N.D. Ind. 2015). In *Messenger* the expert attested to have followed the American Medical Association's standards for work related injuries, but admitted in his deposition to failing to adhere to them without an apparent justification. Short of that sort of "smoking gun" judges are generally willing to admit medical testimony where the expert asserts they engaged in a differential diagnosis, with any missing factors reserved for cross-examination rather than serving as a basis for rejection.

Contrast this largely deferential standard with the legally constructed standard for toxicology testimony to be relevant in many courts after the Ninth Circuit ruling on remand in the *Daubert* dispute. There, for evidence to establish a substance was more likely than not the cause of harm, toxicologists were regularly required to show a doubling of risk (Beyea and Berger 2001: 348-360; Beecher-Monas 2011: 41-43). As critics noted, this had the effect of sharply reducing the use of otherwise valid science in court. Here, by comparison, the same sort of legal construction may be honoring the intent of *Daubert* to make admission the rule, but it may be doing so at the risk of admitting less than rigorous applications of medical science.

This possibility is underscored by the fact that the legal construction of differential diagnosis and differential etiology is not generally what medical science is focused on or how the

field actually uses the terms. As the most recent edition of the *Reference Manual on Scientific Evidence* makes clear, a quest for a specific causal agent such as referenced in *Koho* is not how the term differential diagnosis is actually used in medicine:

In the medical context, by contrast, differential diagnosis refers to a set of diseases that physicians consider as possible causes for symptoms the patient is suffering or signs that the patient exhibits. By identifying the likely potential causes of the patient's disease or condition and weighing the risks and benefits of additional testing or treatment, physicians then try to determine the most appropriate approach—testing, medication, or surgery, for example.

(Wong, Gostin and Cabrera 2011:90-91).

Similarly, actual treating physicians are generally not looking at matters through the lens required by the law. In a case like *Koho*, where the litigant is deceased, “[t]estifying . . . requires making judgments that physicians do not ordinarily make in their profession, making these judgments outside of physicians' customary patient encounters, and adapting the opinion in a way that fits the legal standard.” (Wong, et al. 2011:694). Given that this is so, the socially constructed and widely accepted standards for physician testimony might well be creating a gap between the testimony of medical experts and the actual basis for their expertise.

Moreover, this legally constructed standard allows experts to testify with arguably greater certainty in a legal context than would be supported in their professional roles. As Kuhn (1996: 8-10) postulated, the nature of scientific knowledge, and really knowledge generally, is not a linear accumulation but rather a revolutionary process that contains periods of general consensus (“normal science”) punctuated by shifts in understanding that are rarely accepted at first. The *Reference Manual on Scientific Evidence* reflects the shift that medical decision making for several decades, grappling with prior views about the certainty of medical reasoning and shifting towards recognition of the inherently probabilistic and uncertain nature of many diagnoses. Yet

the legal construct of differential diagnosis and standards such as “more likely than not” or even “reasonable medical certainty” do not track either the actual conduct of physicians or the emerging understanding in the field of inherent uncertainty (Wong, et al. 2011 691-92). Rather than attempt to reconcile this split, multiple cases reviewed cited the legal definition supplied by the *Reference Manual on Scientific Evidence* without acknowledging the authors see a sharp distinction between the conduct of physicians in practice and the practice of physicians in court. This despite the overarching goal of *Daubert* and *Kumho* to ensure that expert testimony is linked to how an expert would analyze similar problems in the context of their non-legal endeavors. Consistent with Jasanoff (1995: 130-131), medical testimony has become a socially constructed standard that may fit the needs of the law but is not an accurate reflection of the state of medical practice. This runs directly counter to the purposes of the *Daubert/Kumho* standards.

In addition to this sort of problem, there is also the underlying risk that an industry or court constructed standard is simply wrong. This risk was made manifest danger in the unusual example of *Cascade Yarns, Inc. v. Knitting Fever, Inc.*, (W.D. Wash. 2012) 2012 U.S. Dist. 15097. In this long running fight over the appropriate labeling of blended fiber wools, Cascade Yarns and others in the industry had investigated their rival’s claims to make a cheaper yet high quality cashmere wool blend through what was apparently industry standard practice. They sent a purported cashmere blend to something called the “Cashmere and Camel Hair Manufacturers Institute” (CCMI), who in turn referred the sample to an eminent bioengineer named Kenneth Langley for “fiber content analysis.” *Cascade Yarns, Inc.*, 2012 U.S. Dist. 15097, *4. Professor Langley used the current versions of quantitative and qualitative testing standards developed aby the American Association of Textile Chemists and Colorists (“AATCC”). The qualitative test,

which purported to permit species level discrimination of fiber content, “includes detailed observation of individual fibers under a microscope both in transverse section and elongated, including observation of the scales on animal fibers.” *Cascade Yarns, Inc.*, 2012 U.S. Dist. 15097, *14. He reported similar results when other industry participants sent him the same wool for sampling, suggesting the test was reliable at least in terms of replicability. In short, if the normal defaults of deferring to industry standards and general acceptance were followed admitting this testimony would be an easy call for Judge Ricardo Martinez.

In this case though, there was a critical problem with relying on Dr. Langley. As it turned out, the CCMI sent reference samples for testing to its membership using the AATCC methods every two years. Those tests showed that before the samples that triggered the litigation were sent to Professor Langley he had performed abysmally in the tests. He could generally (but not always) identify a pure sample of cashmere or yak wool, but frequently misidentified the nature of complex strands. Sometimes he claimed a more complex mix than was actually present, others he saw single species in hybrid mixes. Critically, he had found no cashmere in multiple reference samples with documented cashmere levels similar to the purported makeup of the Knitting Fever samples. In other words, it was entirely possible that the linchpin of the plaintiff’s case – the lack of cashmere fibers – was a baseless accusation.

Moreover, Professor Langley was not alone in his failings – the method itself was unreliable. Forty laboratories participated in the 2007 round of the CCMI’s trials in 2007, all following the AATCC protocols. The results varied wildly in their estimates of blended fibers, both in terms of basic composition and specific makeup of the samples. In fact, “[o]f the 200 tests conducted (5 samples x 40 laboratories), 29 tests were accurate within three percentage

points of the actual value for the fibers (wool, cashmere, yak, and angora).” *Cascade Yarns, Inc.*, 2012 U.S. Dist. 15097, *19. In fact, the errors were so widespread that Professor Langley’s estimates approximated the sample average in several cases, something the plaintiff’s counsel actually tried to use to save his testimony, along with relying on the expert’s long history and general acceptance by his peers. Although Judge Martinez refrained from openly mocking the plaintiff’s counsel, he did archly refer to the former effort to as an attempt at “statistical sleight of hand” and had no problem in rejecting the expert as well as bringing to light the problems of the broader industry.²

While Judge Martinez in this one instance was able to overcome the bias in favor of industry standards and other generally accepted measures, the extraordinary evidence it required raises troubling questions. There is no comparable objective reference for all the practitioners of the general outlines of NFPA 921. As the FJC *Reference Manual on Scientific Evidence* makes clear, the standards of differential diagnosis and differential etiology in court do not track the processes doctors use in treatment. Moreover, there is a growing recognition that diagnosis, as a complex and probabilistic exercise, can have a high degree of error and subjectivity – doctors often make presumptive diagnoses that are ultimately inaccurate, and frequently multiple doctors disagree over how to interpret the same data (Wong, et al. 2011:691-694; Baker 2005: 502-503). This doesn’t mean experts should be rejected as a matter of course, as that would suggest only testimony free from uncertainty should ever be admitted. At the same time, it does suggest more care should be paid to the process by which industry standards or conventions were arrived at

² Interestingly, although not directly on point with the subject of this study, several years later the FBI admitted their own method of microscopic hair analysis was fatally flawed, with erroneous statements in more than 90 percent of cases reviewed by the FBI and the Innocence Project (FBI 2015).

and whether there are any meaningful benchmarks to ensure the methods actually are reliable in the sense of *Daubert/Kumho*.

The *Cascade Yarns* case is also emblematic of why judges likely have such faith in the power of the adversarial system as an element of the admissibility heuristic discussed above. The CCMI's test results were obtained through discovery. They were used effectively to by Knitting Fever's counsel to a degree Judge Martinez was moved to take the extraordinary step of eliminating an otherwise eminently qualified witness and indeed rendering a whole category of evidence suspect. From the perspective of a judge (and likely Knitting Fever), the system worked as designed – the testimony was examined and found wanting on the merits through the tools of adversarial proceedings. But there is no way of knowing how many erroneous conclusions were admitted before this case – nor any certainty as to how widespread this sort of problem truly is in various fields of expertise.

In addition to relying on potentially inaccurate constructions of various methods the judges in this study also relied on the judgment of their peers in prior cases as to particular experts. That is, individual experts with either a positive or negative reputation in prior litigation tend to be treated in the same manner in subsequent cases. This was most notable in MDL cases, such as the *In re Yasmin and Yaz* litigation as well as suits against Novartis and Forest Labs, but it emerged in other contexts as well. Here too, this raises some question as to whether district judges are furthering the intent of the *Daubert/Kumho* trilogy with regards to examining methods rigorously.

Examples of the impact of prior judicial consideration were common in the dataset. As already discussed, in *Dasho* the shot pattern reconstruction technique had been previously ruled

against in state court in a non-precedential case that the district court cited in part as a basis for justifying his own act of exclusion. In *Parker v. Smithfield Packing Co., Inc.* the allowance of competing experts on time work studies was rooted in part on prior admission rulings in other cases. *Parker v. Smithfield Packing Co., Inc* (E.D. N.C. 2010) 2010 U.S. Dist. LEXIS 102441. Similar experts appeared in cases relating to human resources practices, police practices, fire investigation, life planning and damages. Where there was a clear tendency to admit or reject an expert, the same practice was generally followed.

Even where judges claim not to be influenced by prior rulings, it is not always clear how they could possibly avoid it. In *Aurand v. Norfolk Southern Railway Co.*, Judge Philip Simon considered the testimony of Richard Lipsey, a toxicologist. Although he ultimately predicated his rejection on the fact that the plaintiff's report was so conclusory as to create an "analytical gap requir[ing] the exclusion of Dr. Lipsey's testimony" before he got to that issue he noted an oversight in the recitation of the witness' history as a proposed expert. *Aurand v. Norfolk Southern Railway Co* (N.D. Ind. 2011), 802 F.Supp. 2d 950, 958. Specifically, while claiming he had universally been accepted as an expert, he'd actually been rejected in the following colorful terms in Florida district court:

[T]his is the worst example that I have seen in my 15 years experience in the courts as a judge demonstrating what's wrong with expert testimony in our courts of law. It's obvious you can get a Ph.D. to say anything, and this one is prepared to say anything, gratuitous or otherwise...It's just sheerly his opinion -- and I do put that in quotes because, in my opinion, it's absolutely pure and simple, unadulterated speculation, guesswork, just blown-in. There's no scientific basis for any opinion that he has rendered, and I would consider him to be a false expert; and, therefore he will not be permitted in this court, unless the Eleventh Circuit directs otherwise.

Id. 802 F. Supp. 2d at 954 (quoting *Williams v. Orkin Exterminating Co.*, Cause No. 3:95CV30511-LC, DE 275, pp. 55-56).

It is difficult to credit that Judge Simon truly set aside such a prior view after having given it such prominence.

No expert illustrated the impact of prior testimony more clearly, though, than Suzanne Parisian. A former FDA regulator and founder of a consulting business “specializing in matters involving the regulation of United States products by the FDA”, Dr. Parisian was discussed by judges in every court cluster considered. *In re: Yasmin and Yaz*, 2011 U.S. Dist. LEXIS 145593, *47-48. In *Bryant v. Wyeth* Judge Thomas Zilly went through her history both in the context of suits involving hormone replacement therapy and more generally. On at least one subject the only justification offered for ruling the testimony admissible was that other courts had let it in, and the primary driver for deciding to reserve judgment on another subject was the existence of a split in other court’s treatment of the subject. *Bryant v. Wyeth* (W.D. Wash 2012) 2012 U.S. Dist. LEXIS 190912, *11-19. In the Yasmin/Yaz litigation, Judge Herndon used prior endorsements of her testimony as part of the reasoning to reject a challenge to her methods. 2011 U.S. Dist. LEXIS 145593, * 50. In Novartis related litigation in the Western District of North Carolina, Judge Graham Mullen decided to reject a challenge to the entirety of her testimony based on finding that the subject matter of the case was more similar to cases where she’d been admitted than to those in which she had been rejected. *Lemons v. Novartis Pharmaceuticals Corp.*,(W.D. N.C. 2012) 849 F. Supp. 2d 608, 613-14 (Stating “[t]his Court finds it persuasive that all but one Court, *Hogan*, that considered Dr. Parisian’s testimony in an Aredia® or Zometa® case found her testimony to be admissible as a general matter.”) Dr. Parisian thus is a body of law unto herself, distinct from the methods she employs – and with at least one split of authority on whether to admit her.

It is easy to understand why judges would want to be appraised of an expert's prior conduct. Disclosure of prior testimony is one of the requirements of an expert disclosure under Federal Rule of Civil Procedure 26. At the same time, district court rulings are not binding precedent, and in theory a *Daubert/Kumho* assessment should be focused on the application of the expert's methodology in the present case rather than their general credibility. This focus on what other courts think of an expert raises the prospect of the creation of a favored and disfavored caste of experts based primarily on their prior treatment. As in Dr. Parisian's case, the same person can hold both a favored and disfavored status depending on the particular nature of the case under consideration. Countering that impulse is difficult but not impossible – Judge Herndon specifically noted criticisms of Dr. Parisian's manner of testimony or whether she was “uncontrollable and unhelpful” as a witness were irrelevant to his decision at the pre-trial stage. 2011 U.S. Dist. LEXIS 145593, * 52-53. If nothing else, though, the repeated reliance on the opinions of judicial peers underscores that the admission or rejection of a particular expert – especially one with a reputation – may not always rise and fall solely on the consideration their methodology as envisioned by *Daubert*.

Ultimately, it seems that judges cannot always divorce themselves from the human impulse to seek assurance in convention. This reassurance can come from a broad range of sources, from trade associations to authoritative rulings to the non-binding opinions of their peers. Each creates at least the potential for judges to minimize their own efforts in assessing proposed experts. Similar to most such efforts, there is a potential risk involved in such simplifications. In particular, it may lead to a process that is at one legally compliant and yet

ultimately undermines the intent of the *Daubert/Kumho* trilogy to ensure expert testimony reflects the rigor expected of work in a particular field.

C. The Framework Heuristic - Judges' Rulings Demonstrate Significant Cognitive Simplification in Terms of Structure and Substance

As noted above, there has been a suggestion that even though judges in an experimental setting demonstrate the impact of heuristics the act of writing opinions might serve as a limit on their impact (Rachlinski and Wistrich 2017:223). This is certainly a view that appellate judges often take, asserting that in writing they have to consider their colleagues' and broader audience's views as well as their own (Kozinski [1993]2013:116-117). At the same time, Beebe's (2006) study of trademark law suggested that open ended legal standards could lead individual circuits and judges to develop decision heuristics that rendered some factors outcome determinative even as they professed to be weighing the "totality of the circumstances". In the context of *Daubert/Kumho* admissibility determinations, written opinions suggest that judges look for ways to minimize the burden of analyzing the question, both by how they structure the opinion and what they choose to discuss – the framework that is used to assess these types of issues. This in turn suggests that while the goal of liberalizing expert testimony may be met it can come at the expense of close examination of methods.

From the outset, judges signal that they are trying to limit the amount of cognitive effort required to resolve these questions by the structure of the opinions. Judges typically adopted a preferred legal framework and then applied that boilerplate language moving forward. Most prominently in these passages – some as long as several pages, others as short as a single paragraph – would be some recitation of the various elements of the *Daubert/Kumho* line of cases and/or Federal Rule of Evidence 702, often supplemented with circuit authority.

It was in this context that a phenomenon similar to that of Beebe (2006) emerged. While all courts acknowledge the controlling authority, they differ in assessing what this actually means in terms of the process a court must undertake. In its simplest form as commonly expressed in the Fourth Circuit, judges assessed the gatekeeping function as being a matter of assessing only two general factors – reliability and relevance. *See Earp v. Novartis Pharmaceuticals Corp* (E.D. N.C. 2013), 2013 U.S. Dist. LEXIS 129910, *7-8 (“Courts have distilled the requirements of Rule 702 in two crucial inquiries: whether the proposed expert's testimony is relevant and whether it is reliable.”). There is no definitive ordering of these considerations in these cases, and consideration of the expert's qualifications is a component part of reliability. In contrast, the courts in the Ninth Circuit will use the same basic question, but add a threshold requirement of assessing the qualifications of the expert. *Bisson v. BNSF Railway Co* (E.D. Wash 2015), 2015 U.S. Dist. LEXIS 182242. In the Seventh Circuit, by comparison, this same inquiry is often (but not universally) stated as three distinct steps, to be taken in the specific order of qualifications, reliability and relevance. *Sann v. Mastrian* (S.D. Ind. 2012), 2012 U.S. Dist. LEXIS 9107; cf. *In re Yasmin and Yaz (Drospirenone) Marketing, Sales Practices and Products Liability Litigation* (SD. Ill. 2011), 2011 U.S. Dist. LEXIS 145593, *7-8 fn. 3 (noting the different formulations, but concluding “*Chapman* simply combines the first two steps described in *Ervin* as a single test of reliability, whether the analysis is described as a three-step or two-step process does not substantively change the Court's analysis.”)

The primary impact of the framework is to vastly simplify the task in front of the judge. The value of boiler plate as a labor-saving device is shown by how they were re-used both among judges in the same district and by the same judge in different cases. For example, Judges James Dever and Louise Flanagan in the Eastern District of North Carolina used some of the exact same phrasing in their handling of cases, beyond simply applying the same body of precedent. *See Earp v. Novartis Pharmaceuticals Corp* (E.D. N.C. 2013), 2013 U.S. Dist.

LEXIS 129910, *7-8; cf. *SMD Software, Inc. v. Emove, Inc* (E.D. N.C. 2013), 945 F. Supp. 2d 628 at 636-37. In some extreme cases, a judge's application of boilerplate extended beyond the recitation of a legal framework to the substantive analyses of issues. In both *C.A. v. AMLI at Riverbend, L.P.* and *U.S. Automatic Sprinkler, Co., v. The Reliable Automatic Sprinkler Co.*, for example, Judge Sarah Barker of the Southern District of Indiana used the exact same phrasing to dismiss an objection, clearly recycling her effort from the prior case to the latter. See *C.A. v. AMLI at Riverbend, L.P.*, (S.D. Ind. 2010) 2010 U.S. Dist. LEXIS 86530, *10-11; *U.S. Automatic Sprinkler* (S.D. Ind. 2010) 2010 U.S. Dist. LEXIS 29456, *7.

In addition to generally reducing the cognitive burden of writing each decision, the application of the general framework revealed that judges seem to favor the familiar tasks of assessing the qualifications and relevance of witnesses. Assessing witnesses generally in terms of their competence and background is something judges do regularly as part of making credibility determinations. Similarly, relevance and the proper scope of testimony are questions that judges rule on in every case and rooted in well-established legal principles all lawyers know. By comparison, assessing methodology is unique to *Daubert/Kumho* proceedings. This last category requires far more cognitive effort to engage directly and meaningfully. This is the type of situation in which the bounded rationality findings of Simon (1955), the work of Kahneman (2011) on heuristics and biases and the insights into juror behavior of Cooper and Neuhaus (2000) all suggest people will tend take action where they feel confident in their expertise and tend to find a way to avoid direct engagement with the more complex task.

Assessing the qualifications of experts is very similar to assessing the capacities of witnesses in a variety of other circumstances, including competency and credibility. Out of the

three general types of assessment judges frequently cited in their legal frameworks, they actually looked at credentials in the vast majority of cases (over 190 separate entries). Even when the expert's qualifications were not challenged, the opinion frequently recited them at some length. This arguably might be a result of conditioning to preserve the record on appeal – but it also is suggestive of a judge who is otherwise unfamiliar with the subject of testimony being able to find their footing in the expert's credentials.

However, merely possessing stellar academic or experiential qualifications is not a pass for a witness. If anything a misfit between qualifications and proposed testimony seemed to be most likely when the expert was highly qualified- just not for what they were trying to say. For example, two separate experts - a distinguished architect in one case, a workplace safety and human factors design expert in the second - were excluded from testifying as to an alleged design defect in a parking lot for the simple reason that the attorneys offering them did not even bother to point to any direct knowledge of traffic design on their part. *Estate of Myers v. Wal-Mart Stores, Inc* (E.D. N.C. 2011), 2011 U.S. Dist LEXIS 39164; *Roberts v. Menard, Inc* (N.D. Ind. 2011), 2011 U.S. Dist LEXIS 44628. The key is not an absolute weighing of credentials but a relative comparison of credentials to subject matter – such as approving a fire inspector who never took a university course in engineering and did not acknowledge the industry standard for fire investigation but had been in the business of investigating dryer fires for a number of years and followed the general outlines of the standard. *State Farm Fire & Casualty Ins. Co. v. Electrolux Home Products* (N.D. Ind. 2012) 2012 U.S. Dist. LEXIS 188434.

Assessing relevance is another area that judges confront in every case. Finding a lack of relevance is possible, but generally only if the issue the testimony was meant to address has been

resolved or eliminated in some way. It can also occur when the information the expert would convey – such as compliance with a legal standard that is fundamentally different than the matter at issue in the current case – is seen as lacking any connection to the ultimate issue in dispute. Barring this sort of situation, judges can and do summarily find the testimony relevant.

The remaining area examined under the framework is the reliability of the expert's testimony. Nominally, this is at the heart of *Daubert/Kumho*, with a goal of ensuring that expert testimony reflects intellectual rigor consistent with the way an expert would behave outside of the context of litigation. Yet most of these cases have almost no discussion of methods. Moreover, although *Daubert* is frequently taught in law schools for the four factors it set out for evaluating scientific testimony - testability/replicability, error rate, peer review/publication and general acceptance – they feature far less prominently in practice. Only 26 cases were coded as containing a recitation of the four core *Daubert* factors in the substantive portion of the challenge, and even fewer actually analyzed these as factors as a group.

This is in part tied to how *Daubert/Kumho* framed the utility of the factors. Somewhat like the factors in trademark studied by Beebe (2006), the “*Daubert* factors” were expressly couched as nonexclusive list. In expanding the gatekeeping function to all testimony, rather than what the courts deemed “scientific” testimony, *Kumho* noted the flexibility inherent in deciding which factors were relevant would be linked to the facts of a particular case. Essentially, *Daubert*'s specific framework was cabined by *Kumho* to “scientific” testimony, with how to evaluate other testimony left open to development.

However, in a striking difference with trademark law and other fields, courts have generally not sought to augment the list of factors. Rather, it is more typical for both appellate and district courts to find ways to eschew the factors altogether. This is typically achieved by

drawing a line between scientific and non-scientific testimony, despite the intent of *Kumho* to ensure a similar form of gatekeeping is applied in all cases of expert testimony. This divide in turn allows the judge to assess methodology against essentially any standard – which in turn may be something that appears like no standard at all.

Two cases that are emblematic of this are *Bisson* and *Parker*. In *Bisson*, the plaintiff alleged that he had suffered a work related injury in part because of the condition of his work truck. He tendered an expert who concluded, without much discussion of methods, that there were a number of problems with the truck that placed it in a condition to cause this injury when the truck hit a pothole. Yet when the defense pointed out that these conclusions were not the product of a method that would pass muster under the *Daubert* factors, Judge Robert Whaley brushed aside the challenge:

Defendant argues that the methodology used by Mr. Curtis does not meet Fed. R. Evid. 702 nor the factors outlined in *Daubert* to help courts determine reliability. This argument, however, misses the point that not all expert testimony is based on scientific foundation. See *Kumho Tire*, 526 U.S. at 150 ("Engineering testimony rests upon scientific foundations, the reliability of which will be at issue in some cases In other cases, the relevant reliability concerns may focus upon personal knowledge or experience.")

Here, the Court finds that reliability in this instance depends on the expert's knowledge and experience more so than the methodology or theory behind any science. See *United States v. Hankey*, 203 F.3d 1160, 1169 (9th Cir. 2000) ("The *Daubert* factors (peer review, publication, potential error rate, etc.) simply are not applicable to this kind of testimony, whose reliability depends heavily on the knowledge and experience of the expert, rather than the methodology or theory behind it.").

2015 U.S. Dist. LEXIS 182242, *7-8.

In other words, sometimes so long as the expert is testifying to things within their expertise, there would be no way to exclude the testimony. If that were taken to its extremes, it

would be hard to see how any non-scientific testimony, as that term is defined by courts, would ever be subjected to much scrutiny.

Parker reflects a similar conclusion albeit phrased differently. In this case a putative class of employees asserted Fair Labor Standards Act violations linked to their employer's refusal to pay for time spent on various work-related tasks. Their expert's primary method was observational and involved simple measurements of particular tasks to arrive at an average time for each task. The defendant's asserted that the expert's non-statistical method was non-replicable and violated prevailing industry standards for standard setting. Magistrate Judge James Gates rejected this assertion in colorful terms:

In advancing many of their arguments, defendants advocate a rigid application of the factors bearing on the reliability of a methodology listed by the Fourth Circuit in *Tunnell v. Ford Motor Co.*, 245 Fed. Appx. 283, 287 (4th Cir. 2007), namely, error rate, controls on application, peer review, and general acceptance. But the Supreme Court and the Fourth Circuit itself have instructed that the determination on reliability should be a flexible one. "In making its initial determination of whether proffered testimony is sufficiently reliable, the court has broad latitude to consider whatever factors bearing on validity that the court finds to be useful; the particular factors will depend upon the unique circumstances of the expert testimony involved." *Westberry*, 178 F.3d at 261 (citing *Kumho Tire*, 526 U.S. at 150-52). The circumstances here, including the nature of the subject matter as the donning and doffing of equipment as opposed to some intangible and esoteric biological or chemical process, do not demand the rigorous level of scientific precision for admissibility that defendants advocate.

2010 U.S. Dist. LEXIS 102441, *18-19.

Again, the impact of this read broadly is to give lower courts carte blanche to use or refrain from applying the *Daubert* factors in any setting not fitting that particular judge's concept of scientific testing. That in turn allows judges to tremendously reduce the cognitive burdens of evaluating a *Daubert/Kumho* issue. In particular, they can take action that they can reconcile

with legal doctrine without having to engage with the complexities of understanding methodological standards, much less applying them in a broad range of different fields.

Ultimately, the simplified framework allows judges both to sort the contentions of the parties and to rapidly evaluate them. Analyses of qualifications are simple, and generally result in a finding of a reasonable fit of the expert and subject matter of testimony. Analyses of relevance also default to finding the testimony has value outside of extreme cases. Methods, which should logically be the central focus of a *Daubert/Kumho* proceeding, are often reduced to a simple comparison to standards in the field in keeping with the acceptance heuristic and most technical questions brushed aside or reserved for the adversarial process at trial. While it thus significantly simplifies the task of addressing these questions, this heuristic again raises the potential that judges are not always ensuring that the key goal of the *Daubert/Kumho* line of cases – of ensuring that judges serve as meaningful gatekeepers of expert testimony – is consistently met.

IV. Conclusion

This study consistently identified ways in which judges are attempting to lessen their cognitive burdens. They default to admitting experts in all but the most extreme cases and trust in the adversarial process to resolve even complex debates over expert methodology. They accept most testimony so long as it can be shown to be within the mainstream of either an industry's standard such as NFPA 921 or a legal construct such as differential diagnosis. They also accept most testimony when the expert can demonstrate a prior track record of admission. They significantly simplify the task of making these assessments by how they structure their analyses into three primary considerations – qualifications, relevance and methods. Barring

extreme problems, most issues are reserved for cross-examination and they jury's assessment of the weight of the evidence, rather than its admissibility.

All of these simplifications are consistent with findings in other areas of law and teachings from psychology and behavioral economics. District judges are confronted with repeated decisions involving uncertainty under time constraints. It is only logical that they would seek to minimize the cognitive burden of their tasks. Moreover, these simplifications do not seem to lead to rulings that are non-compliant with legal norms and doctrine. If anything, they both meet the exhortations of the Supreme Court to liberalize admissibility while honoring the legal norms of triadic dispute resolution.

It should also be stressed that these are not indications that judges are shirking or attempting to evade their jobs. District court judges have staggering workloads. They are under pressure from both parties and other actors in the court system – up to the Chief Justice of the Supreme Court – to engage and resolve cases expeditiously. This is at most another reminder that judges are human, as Jerome Frank (1936) tried to teach us all.

Moreover, some caution is warranted in light of the size and scope of this study. It only represents the conduct of a minority of judges, and only in this somewhat unique intersection of law and science. This study did not look to quantify the impact of heuristics in terms of potentially improper admissions or sub-optimal trial outcomes. In other more general areas the admissibility heuristic likely disappears. The acceptance heuristic is likely less common, although still quite possible with “repeat player” witnesses such as police officers and forensic technicians in criminal cases as well as other similar situations. The framework heuristic, while likely endemic to all situations in which judges regularly have to issue written rulings, may be

less problematic when judges are dealing with core functions with less of an interface with outside questions of technology, science and expertise.

Nonetheless, the existence and persistence of these heuristics across courts and regions suggest that they could lead to sub-optimal results both because they permit arguably improper testimony to go forward in cases and/or permit “battles of the experts” that ultimately confuse jurors. The *Cascade Yarn* case highlights how an entire field can go years without having its problems exposed, and we cannot know how many other fields have similar issues. Empirical studies have shown multiple experts do not clarify errors but rather tend to cause jurors to reject all the experts. Eliminating the former would require much more effort to establish criteria for evaluating fields of expertise along the lines of the FJC *Reference Manual on Scientific Evidence* as well as requiring judges that are capable of even meaningfully evaluating core methodology questions. The latter would likely require fundamentally shifting legal norms to encourage judges to engage rather than avoid taking a position on issues related to expert testimony, breaking with some of the core traditions of the adversarial system. None of this is likely to be forthcoming in the near term.

Rather, this is a moment that calls for thinking about how we can train lawyers and ultimately judges to be better consumers and advocates of positions rooted in scientific and technical testimony. Perhaps, as Geyh (2016) argues, we need to examine shifting the “legal paradigm” to one that is more in line with other fields of knowledge. Perhaps we should simply admit that such simplifications are inevitable and find ways to minimize their potential negative impact through checklists or counter-heuristics such as the one proposed by Beecher-Monas (2001). At an absolute minimum, though, more work is necessary to understand the impact of

cognitive simplifications on judicial conduct both in immediate decision making and the process of constructing written decisions – because it is undeniably a factor in the outcome of many cases.

References

Books, articles and statutes

- Bainbridge, Stephen M., and Gulati, Mitu. 2002."How Do Judges Maximize? (The Same Way Everybody Else Does-Boundedly): Rules of Thumb in Securities Fraud Opinions."
Emory Law Journal, 51:83-151.
- Baker, Tom. 2005.“Reconsidering the Harvard Medical Practice Study Conclusions about the Validity of Medical Malpractice Claims.” *The Journal of Law, Medicine and Ethics*, 33:501-514.
- Baum, Lawrence. 2006. *Judges and Their Audiences: A Perspective on Judicial Behavior*. Princeton, NJ: Princeton University Press.
- Baum, Lawrence. 2010. “Motivation and Judicial Behavior: Expanding the Scope of Inquiry.” In *The Psychology of Judicial Decision Making*, ed. David E. Klein, and Gregory Mitchell, 1-26. Oxford, UK: Oxford University Press.
- Beebe, Barton. 2006. “An Empirical Study of the Multifactor Tests for Trademark Infringement.” *California Law Review*, 94:1581-1633.
- Beecher-Monas, Erica. 2001. “The Heuristics of Intellectual Due Process: A Primer for Triers of Science.” *NYU Law Review*, 75:1564-1657
- Beecher-Monas, Erica. 2011. “Expert Testimony in Civil Cases.” In *The Future of Evidence: How Science and Technology Will Change the Practice of Law*, ed. Carol Henderson and Jules Epstein, 33-76. Chicago: American Bar Association.
- Beyea, Jan, and Berger Daniel. 2001. "Scientific Misconceptions among Daubert Gatekeepers: The Need for Reform of Expert Review Procedures." *Law and Contemporary Problems*, 64(2/3):327-372.
- Boyd, Christina L. (2013) “She’ll Settle It?” *Journal of Law and Courts*, 1(2):193-219
- Braun, Virginia and Clarke, Victoria. 2006. “Using Thematic Analysis in Psychology.” *Qualitative Research in Psychology* 3:77-101.
- Brekke, Nancy J.; Enko, Peter J.; Clavet, Gail; and Seelau, Eric. 1991. "Of Juries and Court-Appointed Experts: The Impact of Nonadversarial versus Adversarial Expert Testimony." *Law and Human Behavior*, 15:451-475.
- Buchman, Jeremy. 2007. “The Effects of Ideology on Federal District Judges’ Decisions to Admit Expert Testimony.” *American Politics Research*, 35:671-693.
- Cecil, Joe S. and Willging, Thomas E. 1993. “Court-Appointed Experts: Defining the Role of Experts Appointed Under Federal Rule of Evidence 706.” Federal Judicial Center.
- Charmaz, Kathy. 2006. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. Thousand Oaks: Sage Publications.

- Cooper, Joel and Neuhaus, Isaac M. 2000. The "Hired Gun" Effect: Assessing the Effect of Pay, Frequency of Testifying, and Credentials on the Perception of Expert Testimony. *Law and Human Behavior*, 24:149-171.
- Dixon, Lloyd and Gill, Brian. 2002. "Changes in the Standards for Admitting Expert Evidence in Federal Civil Cases Since the *Daubert* Decision." *Psychology, Public Policy, and Law* 8:251–308.
- Epstein Lee. 2016. "Some Thoughts on the Study of Judicial Behavior." *Wm. & Mary Law Rev.* 57(6):2017-73.
- Epstein, Lee, and Knight, Jack. 2013. "Reconsidering judicial preferences." *Annual Review of Political Science* 16:11-31.
- Federal Bureau of Investigation. 2015. "FBI Testimony on Microscopic Hair Analysis Contained Errors in at Least 90 Percent of Cases in Ongoing Review" Press Release. Available at <https://www.fbi.gov/news/pressrel/press-releases/fbi-testimony-on-microscopic-hair-analysis-contained-errors-in-at-least-90-percent-of-cases-in-ongoing-review> accessed on January 16, 2018.
- Federal Judicial Center, National Research Council of the National Academies. 3rd ed. 2011. *Reference Manual on Scientific Evidence*. Washington, D.C.: The National Academies Press.
- Frank, Jerome. 4th printing 1936. *Law and the Modern Mind*. New York: Coward-McCann.
- Gatowski, Sophia I.; Dobbin, Shirley A.; Richardson, James T.; Ginsburg, Gerald P.; Merlino, Mara L.; and Dahir, Veronica. 2001. "Asking the Gatekeepers: A National Survey of Judges on Judging Expert Evidence in a Post-*Daubert* World." *Journal of Law and Human Behavior* 25:433–58.
- Geyh, Charles Gardner. 2016. *Courting Peril: The Political Transformation of the American Judiciary*. New York, NY Oxford University Press.
- Guthrie, Chris; Rachlinski, Jeffrey J.; and Wistrich, Andrew J. 2007. "Blinking on the Bench: How Judges Decide Cases." *Cornell Law Review*, 93:1-43.
- Helland, Eric and Klick, Jonathan. 2012. "Does Anyone Get Stopped at the Gate? An Empirical Assessment of the *Daubert* Trilogy in the States." *Supreme Court Economic Review*, 20:1-33.
- Hornby, D. Brock. [2007] 2013. "The Business of US District Courts." In *Judges on Judging*, 4th ed., ed. David M. O'Brien, 106-114. Thousand Oaks, CA: CQ Press.
- Hsieh, Hsiu-Fang and Shannon, Sarah E. 2005. "Three Approaches to Qualitative Content Analysis." *Qualitative Health Research*, 15:1277-88.
- Jasanoff, Sheila. 1995. *Science at the Bar: Law, Science, and Technology in America*. Cambridge, MA: Harvard University Press.
- Jasanoff, Sheila. 2008. "Representation and Re-Presentation in Litigation Science." *Environmental Health Perspectives*, 116:123-129.

- Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.
- Kahneman, Daniel and Tversky, Amos. 1979. "Prospect Theory: An Analysis of Decision Under Risk." *Econometrica*, 47(2): 263-291.
- Keck, Thomas. 2014. *Judicial Politics in Polarized Times*. Chicago: University of Chicago Press.
- Kozinski, Alex. [1993] 2013. "What I ate for Breakfast and Other Mysteries of Judicial Decision Making." In *Judges on Judging*, 4th ed., ed. David M. O'Brien, 115-120. Thousand Oaks, CA: CQ Press.
- Kuhn, Thomas. 3rd ed. 1996. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Levett, Lora M. and Kovera, Margaret B. 2008. "The Effectiveness of Opposing Expert Witnesses for Educating Jurors About Unreliable Expert Evidence." *Law and Human Behavior*, 32:363-374.
- Levitin, Daniel J. 2017. *Weaponized Lies- How to Think Critically in the Post-Truth Era*. New York: Dutton.
- McAuliff, Bradley D. and Duckworth, Tejah D. 2010. "I Spy With My Little Eye: Jurors' Detection of Internal Validity Threats in Expert Evidence." *Law and Human Behavior*, 34:489-500.
- McAuliff, Bradley D., Kovera, Margaret B. and Nunez, Gabriel. 2009. "Can Jurors Recognize Missing Control Groups, Confounds, and Experimenter Bias in Psychological Science? *Law and Human Behavior*, 33:247-257.
- Merlino, Mara L., Springer, Victoria and Sigillo, Alexandra E. (2011). "The Social Construction of the Admissibility of the Most Frequently Proffered Varieties of Expert Testimony." In *The Future of Evidence: How Science and Technology Will Change the Practice of Law*, ed. Carol Henderson and Jules Epstein, 1-32. Chicago: American Bar Association.
- Rachlinski, Jeffrey and Wistrich, Andrew. 2017. "Judging the Judiciary by the Numbers: Empirical Research on Judges." *Ann. Rev. Law Soc. Sci.* 13:203-229.
- Rocco, Phillip. 2017. "Justice Roberts said political science is 'sociological gobbledegook.' Here's why he said it, and here's why he's mistaken." *The Washington Post*. Available at www.washingtonpost.com/news/monkey-cage/wp/2017/10/04/justice-roberts-said-political-science-is-sociological-gobbledegook-heres-why-he-said-it-and-why-hes-mistaken/?utm_term=.96ab8de54bfd Accessed Nov. 13, 2017.
- Schauer, Frederick, and. Spellman, Barbara A. 2013. "Is Expert Evidence Really Different?" *Notre Dame L. Rev.*, 89:1-26.
- Seawright, Jason and Gerring, John. 2008. "Case Selection Techniques in Case Study Research: A Menu of Qualitative and Quantitative Options." *Political Research Quarterly*, 61:294-308.

- Sevier, Justin. 2014. "The Truth-Justice Tradeoff: Perceptions of Decisional Accuracy and Procedural Justice in Adversarial and Inquisitorial Legal Systems." *Psychology, Public Policy, and Law*, 20:212-224.
- Simon, Herbert. 1955. "A Behavioral Model of Rational Choice". *Quarterly Journal of Economics*, 69: 99-118.
- Thaler, Richard H. and Sunstein, Cass R. 2008. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New Haven, CT: Yale University Press.
- Tversky, Amos and Kahneman, Daniel. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science*, 185:1124-1131.
- Watson, Randy. 2015. "Is it Arson? The NFPA 921 Effect". *In-House Defense Quarterly*, Spring 2015:24-27, 84.

Wong, John B., Gostin, Lawrence O. and Cabrera, Oscar A. 2011. "Reference Guide on Medical Testimony." In *Reference Manual on Scientific Evidence*, 3rd ed., Federal Judicial Center, National Research Council of the National Academies, 687-746. Washington, D.C.: The National Academies Press.

Cases

3600 Michigan Co., Ltd. v. Infra-Metals, Co. N.D. Ill. 2011. 2011 U.S. Dist. LEXIS 574.

Agrigenetics v. Pioneer Hi-Bred International, Inc. S.D. Ind. 2010. 2010 U.S. Dist. LEXIS 12003.

Apilad v. North American Gay Amateur Athlete Alliance. W.D. Wash. 2011. 2011 U.S. Dist. LEXIS 159575.

Aurand v. Norfolk Southern Railway Co. N.D. Ind. 2011. 802 F.Supp. 2d 950.

Bisson v. BNSF Railway Co. E.D. Wash. 2015. 2015 U.S. Dist. LEXIS 182242.

Bryant v. Wyeth. W.D. Wash. 2012. 2012 U.S. Dist. LEXIS 190912.

C.A. v. AMLI at Riverbend, L.P. S.D. Ind. 2010. 2010 U.S. Dist. LEXIS 86530.

Cascade Yarns, Inc. v. Knitting Fever, Inc. W.D. Wash. 2012. 2012 U.S. Dist. 15097.

Citizens Ins. Co. of the Midwest v. LG Electronics USA, Inc. S.D. Ind. 2012. 2012 U.S. Dist. LEXIS 1127128.

Dasho v. City of Federal Way. W.D. Wash. 2015. 101 F. Supp. 3d 1025.

Daubert v. Merrill Dow Pharmaceuticals, Inc. 1993. 509 U.S. 579.

Delarosa v. Speedway, LLC. S.D. Ill. 2013. 2013 U.S. Dist. LEXIS 125104.

Earp v. Novartis Pharmaceuticals Corp. E.D. N.C. 2013. 2013 U.S. Dist. LEXIS 129910.

Estate of Myers v. Wal-Mart Stores, Inc. E.D. N.C. 2011. 2011 U.S. Dist LEXIS 39164.

Fuentes v. Miller. N.D. Ind. 2015. 2015 U.S. Dist. LEXIS 82978.

General Electric Co. v. Joiner. 1997. 522 U.S. 136.

Koho v. Forest Labs, Inc. W.D. Wash 2015. 2015 U.S. Dist. LEXIS 180860.

Koho v. Forest Labs, Inc. W.D. Wash 2015. 2015 U.S. Dist. LEXIS 46050.

Kumho Tire Co. v. Carmichael. 1999. 526 U.S. 137.

Lemons v. Novartis Pharmaceuticals Corp. W.D. N.C. 2012. 849 F.Supp. 2d 608.

McClellan v. I-Flow Corp. D. Ore. 2010. 710 F. Supp. 2d 1092.

Oswalt v. Resolute Industries, Inc. W.D. Wash 2012. 2012 U.S. Dist. LEXIS 190123.

Parker v. Smithfield Packing Co., Inc. E.D. N.C. 2010. 2010 U.S. Dist. LEXIS 102441.

Roberts v. Menard, Inc. N.D. Ind. 2011. 2011 U.S. Dist LEXIS 44628.

Ruppel v. Kucanin. N.D. Ind. 2011. 2011 U.S. Dist. LEXIS 167503.

Sann v. Mastrian. S.D. Ind. 2012. 2012 U.S. Dist. LEXIS 9107.

Severn Peanut Co., Inc. v. Industrial Fumigant Co. E.D. N.C. 2014. 2014 U.S. Dist. LEXIS 34507.

SMD Software, Inc. v. Emovve, Inc. E.D. N.C. 2013. 945 F. Supp. 2d 628.

Stachon v. Woodward. N.D. Ind. 2015. 2015 U.S. Dist. LEXIS 129958.

State Farm Fire & Casualty Ins. Co. v. Electrolux Home Products. N.D. Ind. 2012. 2012 U.S. Dist. LEXIS 188434.

United States v. Johnson. M.D. N.C. 122 F. Supp. 3d 272.

U.S. Automatic Sprinkler, Co., v. The Reliable Automatic Sprinkler Co. S.D. Ind. 2010. 2010 U.S. Dist. LEXIS 29456.

In re Yasmin and Yaz (Drospirenone) Marketing, Sales Practices and Products Liability Litigation. S.D. Ill. 2011. 2011 U.S. Dist. LEXIS 145552.

In re Yasmin and Yaz (Drospirenone) Marketing, Sales Practices and Products Liability Litigation. S.D. Ill. 2011. 2011 U.S. Dist. LEXIS 145593.

In re Yasmin and Yaz (Drospirenone) Marketing, Sales Practices and Products Liability Litigation. S.D. Ill. 2011. 2011 U.S. Dist. LEXIS 145989.