

Florida State University College of Law Scholarship Repository

Scholarly Publications

2015

Signal vs. Noise: Some Comments on Professor Stein's Theory of Evidential Efficiency

Mark Spottswood

Florida State University College of Law

Follow this and additional works at: <http://ir.law.fsu.edu/articles>



Part of the [Evidence Commons](#)

Recommended Citation

Mark Spottswood, *Signal vs. Noise: Some Comments on Professor Stein's Theory of Evidential Efficiency*, 66 *ALA. L. REV.* 471 (2015),
Available at: <http://ir.law.fsu.edu/articles/103>

This Article is brought to you for free and open access by Scholarship Repository. It has been accepted for inclusion in Scholarly Publications by an authorized administrator of Scholarship Repository. For more information, please contact bkaplan@law.fsu.edu.

HEINONLINE

Citation: 66 Ala. L. Rev. 471 2014-2015



Content downloaded/printed from
HeinOnline (<http://heinonline.org>)
Fri May 15 13:15:57 2015

-- Your use of this HeinOnline PDF indicates your acceptance of HeinOnline's Terms and Conditions of the license agreement available at <http://heinonline.org/HOL/License>

-- The search text of this PDF is generated from uncorrected OCR text.

-- To obtain permission to use this article beyond the scope of your HeinOnline license, please use:

[https://www.copyright.com/ccc/basicSearch.do?
&operation=go&searchType=0
&lastSearch=simple&all=on&titleOrStdNo=0002-4279](https://www.copyright.com/ccc/basicSearch.do?&operation=go&searchType=0&lastSearch=simple&all=on&titleOrStdNo=0002-4279)

SIGNAL VS. NOISE: SOME COMMENTS ON PROFESSOR STEIN'S THEORY OF EVIDENTIAL EFFICIENCY

*Mark Spottswood**

ABSTRACT

In this Essay, I examine Professor Stein's intriguing new theory of evidential efficiency, which posits that judges should admit evidence whenever it has a sufficiently high "signal-to-noise" ratio." I explore a slightly different definition of the concepts of "signal" and "noise" than Stein, based upon likelihood ratio values rather than the underlying probabilities of events, and I explain why these altered concepts may be analytically superior. Additionally, I call into question the strength of the connection between the signal-to-noise ratio of a piece of evidence and the costs of admitting it at trial. Nevertheless, Stein's project is worthy of great praise because it focuses our attention on the fact that evidentiary rules have many costs beyond their direct contributions to outcome accuracy. Failing to consider these costs does great harm to individual litigants, the justice system, and society at large.

| | |
|--|-----|
| ABSTRACT..... | 471 |
| INTRODUCTION | 471 |
| I. WHAT IS THE SIGNAL, AND WHAT IS THE NOISE? | 473 |
| II. THE SIGNAL-TO-NOISE RATIO AND EVIDENTIAL EFFICIENCY ARE NOT EQUIVALENT CONCEPTS | 483 |
| III. WHY THE TURN TOWARDS EFFICIENCY ANALYSIS IS ESSENTIAL | 490 |
| CONCLUSION..... | 492 |

INTRODUCTION

Evidence is an unusual field of legal study, in part because so many of its devotees doubt, from time to time, that it should exist at all. Jeremy Bentham, the first great scholar of proof, devoted his life to abolishing all restrictions on the admissibility of evidence, seeking a regime of "free proof."¹ Two simple intuitions explain the enduring appeal of his

* Assistant Professor, Florida State University College of Law. I am grateful to Alex Stein and Murat Mungan for illuminating exchanges on this topic.

viewpoint. First, even if some evidence is misleading, we can communicate to one another the *reasons* why it is misleading, at which point it is much harder to explain why it creates any danger of error. Second, even weak evidence is not valueless, and given the high costs of error at trial, it seems obvious that we should consider every possible item of information that might help us arrive at an accurate decision. Put these two facts together, and the appeal of a system in which the fact finder is permitted to hear all possible relevant evidence is hard to resist. Bentham was quite successful in his campaign to eliminate the proof restrictions that prevailed in his day,² but since then we have replaced the rules that Bentham attacked with a new set of rules that is equally, if not more, complex. For those who agreed with his original arguments, it can seem hard to justify this state of affairs.

Alex Stein's new article, *Inefficient Evidence*, explores this tension from a new and welcome direction.³ His central claim is that even if a piece of evidence is not misleading, it may be costly, and those high costs accumulate in the aggregate, especially in jurisdictions with high levels of litigation.⁴ As a result, it may be possible to justify many proof restrictions via cost-benefit analysis, even if they exclude evidence that provides a modest informational benefit.⁵ He bravely attempts to defend a number of controversial and complicated doctrines—most notably the modern American hearsay rule and its exceptions—on cost-benefit grounds.⁶ Even if this was all he had done, his article would be worthy of careful study.

But beyond these basic policy questions, Stein also offers us an intriguing new theory of admissibility. He borrows the concept of the “signal-to-noise ratio,” or SNR, from the fields of communication science and engineering,⁷ and attempts to link the concept of evidential efficiency to a repurposed version of this formula.⁸ In order to do so, he has to define what “signal” and “noise” refer to in the world of evidence, and he must also show why high SNR evidence is efficient and why low SNR evidence is not. Put together, this is a project of grand scope, which attempts to give us a mathematical technique that both describes and justifies the existing landscape of American evidence law.

1. See generally WILLIAM L. TWINING, THEORIES OF EVIDENCE: BENTHAM AND WIGMORE 3, 66–75 (1985).

2. See David P. Bryden & Roger C. Park, “Other Crimes” Evidence in Sex Offense Cases, 78 MINN. L. REV. 529, 561 (1994).

3. See Alex Stein, *Inefficient Evidence*, 66 ALA. L. REV. 423 (2015).

4. *Id.* at 429–35.

5. *Id.* at 428–29.

6. *Id.* at 443–50.

7. See, e.g., JOHN R. PIERCE & A. MICHAEL NOLL, SIGNALS: THE SCIENCE OF TELECOMMUNICATIONS 29, 51–54 (1990).

8. See Stein, *supra* note 3, at 435–43.

In this response essay, I will explore three questions with respect to Stein's theory and its implementation. In Part I, I will analyze his conception of signal and noise. As I will explain, the way he defines and describes these concepts is problematic, producing some troubling internal inconsistencies. I offer, as a friendly amendment, the suggestion that we might instead define the signal and the noise as functions of subjective likelihood ratios, and explain the advantages of doing so. In Part II, I explore the connection between the signal-to-noise ratio of evidence and the costs of admitting it, and argue that the connection is less direct than Stein has asserted. Finally, in Part III, I explain why Stein's overall project is so important to the future of evidence theory, drawing attention to some additional ways that costly trials undermine our shared ambitions for the American system of justice. If we, as scholars, continue to focus on the accuracy impacts of evidence rules without also attending carefully to their costs, we will fail to offer policy recommendations that are worthy of credence.

I. WHAT IS THE SIGNAL, AND WHAT IS THE NOISE?

At the outset, it is worth spending some time to make sure we understand clearly what Stein means when he speaks of evidential "signal" and evidential "noise." The standard usage of these terms is a way to discuss the clarity with which a mode of communication transmits information from one place to another.⁹ In that usage, the numerator of the ratio is the average power of the "signal," meaning the message that users intend to transmit, over a given channel of communication.¹⁰ The denominator of the ratio is the average power of the omnipresent ambient "noise," which is always present in any communication channel.¹¹ The ratio between these two quantities quantifies the extent to which the message can be clearly discerned over the background noise, and is commonly abbreviated as follows:¹²

$$SNR = \frac{P(\text{signal})}{P(\text{noise})}$$

In order to adapt this formula to be useful for analyzing evidence, Stein must redefine some of its central terms; evidence law, after all, is not

9. *Id.* at 424–25.

10. *Id.* at 425.

11. *Id.*

12. See PIERCE & NOLL, *supra* note 7, at 29.

primarily concerned, with making sure that witnesses can be heard over the air conditioning! To do so, Stein first identifies the concept of “signal” with “information reliable enough to allow the fact finders to determine the probability of [an] underlying allegation,” and “noise” with “the exact opposite,” meaning information that does not allow factfinders to accurately decide the case.¹³ Taken at face value, this first set of definitions might seem to indicate that “signal” and “noise” are a set of binary (and opposite) labels, only one of which could properly be applied to a given piece of evidence. But he soon goes on to reframe the concepts so that each can be assessed with regard to a particular evidence item, by reference to the “probabilities to which a given piece of information gives rise.”¹⁴ By assuming that multiple probabilities flow from a given evidence item, he is then able to equate the signal with the central tendency of those probabilities and the noise with their variance.¹⁵ But before we go so far down the road with him, it is worthwhile to stop and ponder what exactly these multiple probabilities are supposed to represent as a conceptual matter. Referring to “probabilities” without more is not very meaningful; since a probability is a measure of our uncertainty, we must identify what underlying fact or value we are uncertain about. So in this section, I wish to dissect what, exactly, the concepts of signal and noise might be.

The first question, then, is this: When Stein speaks of evidence that “gives rise to probabilities,”¹⁶ what are these probabilities *of*? Stein does not explain this very clearly, so it requires a close look to be sure. At times, he seems to suggest that they are “probabilit[ies] [of] the defendant’s guilt,”¹⁷ but on closer examination that cannot be right. The problem with that interpretation is that Stein equates higher “probabilities” with a higher “signal” in the evidence, and hence, assumes that lower probabilities equate with lower evidential value.¹⁸ Nevertheless, evidence suggesting that the probability of the defendant’s guilt is very low is extremely valuable evidence. In fact, given the frequent assumption that convictions of the innocent are a greater injustice than failures to convict a guilty party, it could easily be argued that evidence suggesting that the probability of guilt is extremely low is among the highest value proof that could exist. As a result, if we follow Stein’s lead on this point, we are left with an untenable possibility: If a signal close to 0 is meant to indicate that evidence is

13. Stein, *supra* note 3, at 424.

14. *Id.* at 425.

15. *Id.*

16. *Id.* at 451.

17. *Id.* at 433, 456 (discussing varying probabilities of guilt as the basis for signal/noise calculations).

18. *Id.* at 425–26.

equally powerful as if it had a signal close to 1, then we face the problem that the SNRs for the two values will be very different, despite the fact that both items have the same evidential strength and the same amount of noise.

For example, assume that item A has a signal of 0.1 and a noise of 0.1, while item B has a signal of 0.9 and a noise of 0.1. If the low signal is meant to prove innocence as strongly as the high signal proves guilt, then the SNR of A and B should be the same; but in fact Stein's formula would give item A an SNR value of 4.5 while only giving B a value of 1. Clearly, then, we cannot treat a high and low signal as equivalent while maintaining the formalism of the SNR ratio.

At other times, Stein seems to be using a narrower conception, such that the probability at issue is the likelihood, given the evidence proffered by a party, that the party's contention about a particular historical fact is true.¹⁹ For instance, when Stein discusses the relation between the SNR and the hearsay rule, he refers not to probabilities of guilt or innocence, but instead to probabilities of the hearsay statement being true.²⁰ Construing the ideas of signal and noise this way avoids the disparities that arise from tying the probability to one party's overall success, but it creates other problems. The problem with this conception is that it fails to connect the concept of "signal" with the *importance* of the evidence to the parties' overall case. Any useful concept of the "signal" in evidence, it seems to me, must track not just the power of the evidence to prove some fact, but also the materiality of that fact to the larger dispute. We should be willing to tolerate higher costs to prove facts that are central to parties' cases, and lower costs to prove peripheral details that are unlikely to change anyone's mind, if we wish to strike an effective balance between the utility and the cost of evidence. By contrast, connecting the signal with the probability that any allegation is true means that we would be assigning the same signal to either crucial or near-irrelevant evidence, so long as the allegation was equally likely to be true based on the evidence proffered. So, although the "probability of the underlying event" is a better conception of the signal than the "probability of guilt," it still leaves something to be desired.

As a friendly amendment, we might posit instead that Stein's "probabilities" do not measure the likelihood of any particular event or proposition, but are instead representations of evidential "strength"—in other words, how effective that item of evidence is in changing our minds about the case. One common way to describe the strength of an individual

19. *Id.* at 424 (referring to the "probability of the underlying allegation"); see also *id.* at 444.

20. *Id.* 444–45.

piece of evidence item is by examining its *likelihood ratio*.²¹ This ratio is expressed as the probability of observing the evidence given the hypothesis of guilt, on the one hand, divided by the probability of observing the evidence given the hypothesis that the defendant is not guilty, on the other:

$$LR(e) = \frac{P(e|G)}{P(e|NG)}$$

Multiple evidence scholars have argued that these ratios provide a useful means of assessing evidential strength,²² and even though they do have their limitations,²³ they have the advantage of giving us a quantification of evidential strength (unlike some competing alternatives) which is a necessity if we are going to use them in conjunction with the signal-to-noise formula. Furthermore, unlike the “probability of the truth of the underlying allegation” formulation discussed above, likelihood ratios respond to variations in both the likelihood that the underlying fact is true, and the materiality of that fact to the dispute.²⁴

But before we plug likelihood ratios into Stein’s formula, we must first transform them into a workable form. Stein’s formula requires an integer measure that ranges from 0 to 1 in both the numerator and the denominator,²⁵ whereas likelihood ratios quantify strength in a very different way. First, a likelihood ratio that is equal to 1 generally indicates that the evidence has minimal value because it is equally likely to arise whether the defendant is innocent or guilty.²⁶ Second, a likelihood ratio that

21. See Edward K. Cheng, *Reconceptualizing the Burden of Proof*, 122 YALE L.J. 1254, 1266–68 (2013) (explaining the methodology of likelihood ratio analysis); Louis Kaplow, *Likelihood Ratio Tests and Legal Decision Rules*, 16 AM. L. & ECON. REV. 1, 5–10 (2014) (same).

22. See Cheng, *supra* note 21, at 1266–68; Kaplow, *supra* note 21, at 5–10; see also DAVID A. SCHUM, *THE EVIDENTIAL FOUNDATIONS OF PROBABILISTIC REASONING* 218–22 (1994); see generally Jonathan J. Koehler, *On Conveying the Probative Value of DNA Evidence: Frequencies, Likelihood Ratios, and Error Rates*, 67 U. COLO. L. REV. 859 (1996).

23. For some discussion on the limits of probabilistic analysis of evidence, see generally PAUL THAGARD, *HOT THOUGHT: MECHANISMS AND APPLICATIONS OF EMOTIONAL COGNITION* 141–45 (2006); Ronald J. Allen & Alex Stein, *Evidence, Probability, and the Burden of Proof*, 55 ARIZ. L. REV. 557 (2013); Ronald J. Allen & Michael S. Pardo, *The Problematic Value of Mathematical Models of Evidence*, 36 J. LEGAL STUD. 107 (2007).

24. See SCHUM, *supra* note 22, at 220–21 (explaining that the likelihood ratio function can be expanded to separately account for the probability that testimonial evidence is true and the subsequent probability that the hypothesis of guilt is true given that probability); see also *id.* at 290–365 (analyzing the many stages involved in moving from evidence to likelihood ratios about events, and in turn from those ratios to inferring the likelihood ratio for the ultimate hypotheses under consideration).

25. See Stein, *supra* note 3, at 425.

26. There are, in fact, important exceptions to this principle. For instance, there will often be items of evidence that do not favor either side in the case, but which provide crucial context, without which the jury could not make sense of either side’s account of the events under dispute. Such evidence

is much greater than 1 implies that evidence of guilt strongly outweighs evidence of innocence, so that evidential strength increases as the likelihood ratio trends towards infinity. Third, when the likelihood ratio gets smaller than 1, the evidence grows more powerful again, but this time it helps to prove innocence rather than guilt. We might summarize these relations with Stein's concept of the evidential signal as follows:

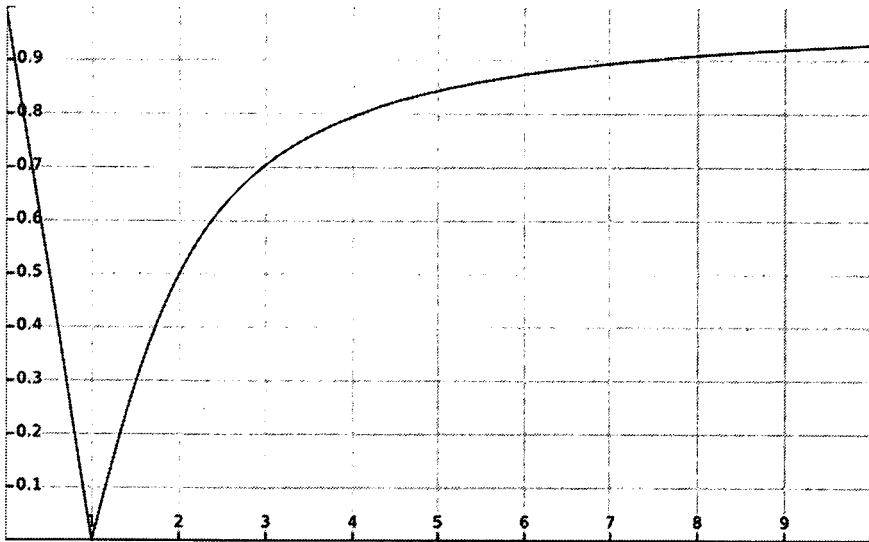
| “Signal” or “S” | Likelihood Ratio or “LR” | Linguistic Description |
|------------------------|---------------------------------|---|
| $S \approx 1$ | LR approaches ∞ | Evidence that implies guilt with maximal strength |
| $S \approx 1$ | $LR \approx 0$ | Evidence that implies innocence with maximal strength |
| $0 < S < 1$ | $1 < LR < \infty$ | Evidence favoring an inference of guilt |
| $0 < S < 1$ | $0 < LR < 1$ | Evidence favoring an inference of innocence |
| $S \approx 0$ | $LR \approx 1$ | Equivocal evidence, favoring neither inference |

Or to make things easier to visualize, the relation between Stein's “signal” and likelihood ratios could also be expressed as a plottable function²⁷ relating likelihood ratio values on the x-axis with signal values on the y-axis:

can, and should, be admitted even though it has a likelihood ratio that is very close to 1. For further discussion on this point, see *infra* Part II.

27. The function that creates this plot is as follows:

$$S(E) = \begin{cases} 1 - LR(E), & \text{for } 0 < LR \leq 1, \\ 0, & \text{for } LR = 1, \\ \frac{2}{\pi} \tan^{-1}(LR(E) - 1), & \text{for } LR > 1 \end{cases}$$



What this discussion implies, I think, is that Stein is wrong to tie his conception of “signal” to the concept of probability. Instead, it seems more workable to say that the signal of evidence is a measure of its power to persuade, which we can derive from the likelihood ratio, but which is not, by itself, any kind of probability. Alternatively, readers who dislike the analytical complexity that likelihood ratios introduce might choose instead

The choice of the inverse tangent function and the applicable parameters were made in order to approximately implement the relations that would be ideal between the likelihood ratio function and the signal function. In particular, it is desirable that $S(E)$ approach 1 as $LR(E)$ approaches infinity, and that the $S(E)$ be approximately equal for each pairing, $LR=1/n$ & $LR=n$, where n is a positive integer. In other words, evidence that makes it twice as likely that the defendant is guilty should have the same signal value as evidence making it half as likely that he is guilty. The function and the corresponding graph do not implement the second constraint perfectly, but they get it approximately right, which should be helpful to at least advance the discussion regarding the concept of evidential “signal.” I am grateful to Murat Mungan for a helpful discussion on these issues.

If we wished to find a quantification of evidential import and strength that would be easier to relate to a “signal” that must vary between 0 and 1, we might alternatively employ likelihood differences, which take the form $LD(e) = P(G|e) - P(NG|e)$. Likelihood differences have the advantage that they value confirming and disconfirming evidence on an identically weighted scale, so that proof favoring guilt grows stronger as LD approaches positive infinity, and proof favoring innocence grows stronger as LD approaches negative infinity. Taking the absolute value, $|LD(e)|$, then gives us a measure of evidentiary signal that is lowest at 0 and increases towards infinity. This avoids the need to specify a relation that is non-differentiable because we can then define $S(e)$ as approximately equal to some function such as $S(e) = \frac{2}{\pi} \tan^{-1} |LD(e)|$. Despite this convenience, I prefer to express the signal as a function of likelihood ratios, given that they are used with far greater frequency both in evidence scholarship and in other fields, and given that the difference measure has some odd properties, such as its tendency to accord lowered weight to evidence whenever the likelihood of its arising is generally low, even if it is highly diagnostic when it arises. Cf. SCHUM, *supra* note 22, at 220 (briefly discussing the likelihood difference approach).

to discard probabilities entirely and consider the signal as a simple scalar measure of evidential “strength” or “power,” without defining those terms probabilistically.

With that working definition of signal, we can then ask the logically subsequent question: “What is noise?” Stein describes “noisy” evidence as that which might allow fact finders to ascribe “virtually any probability to the defendant’s guilt.”²⁸ Later, in his appendix, he describes the relation between signal and noise over a given range of “probabilities” by saying the signal is the average of those probabilities and the noise is the standard deviation of those probabilities, once again suggesting that an underlying “something” is varying.²⁹ Following from my earlier suggestion that signal should be understood as a function of likelihood ratios for each evidence item, we could go further and say that noise is a measure of likelihood ratio *variance*. Or, if we were willing to tolerate the limits of Stein’s suggestion that signal is an expression of the probability that proof proves the immediate underlying event, we might also wish to follow his lead in referring to noise as a measure of the variance in that probability. But that draws us a bit deeper into the rabbit hole, forcing us to answer a subsequent question: “What *kind* of probabilities or likelihood ratios are we working with here?” What we need are measures of uncertainty that vary in a particular way, producing a pattern of potential results for the signal function, of which the central tendency is the best measure of the “true signal.”³⁰ As we shall see, this will be trickier than it sounds.

Most obviously, for it to make sense to talk about likelihood ratio³¹ variance, we cannot define the probabilities that make up the ratio as any kind of objective measure of *epistemic* uncertainty. An epistemic probability is a measure of the “degree of justification” we possess to believe a particular proposition based upon the evidence before us.³² As an

28. See Stein, *supra* note 3, at 433.

29. See *id.*

30. For readers who prefer to avoid likelihood ratios, we might say instead that we are searching for some kind of value that we can represent as a scalar between 0 and 1, that takes multiple values across some distribution, but that retains the signal (defined as a non-probabilistic measure of evidential strength) as its mean. Even with this amendment, it would seem that we still need to introduce some kind of uncertainty into the mix to account for the variations within the distribution.

31. Because I think the likelihood ratio basis for the signal function is more defensible, I will refer to it through the remainder of this discussion when trying to assess what kind of conception of uncertainty we are trying to quantify. Readers who disagree and prefer the original formulation should feel free to mentally substitute Stein’s “probability of underlying events” formulation, because the same concerns arise with respect to either formulation.

32. See Vern R. Walker, *Direct Inference in the Lost Chance Cases: Factfinding Constraints Under Minimal Fairness to Parties*, 23 HOFSTRA L. REV. 247, 259 n.26 (1994); Vern R. Walker, *Preponderance, Probability and Warranted Factfinding*, 62 BROOK. L. REV. 1075, 1080–90 (1996) [hereinafter Walker, *Preponderance*].

objective assessment of how much warrant we have to reach a particular conclusion on the facts before us, an epistemic probability is not meant to describe what any individual actually believes. It is, instead, a normative prescription telling us how much confidence we *should* have, based on our evidence, if we possessed idealized powers of analysis and reasonableness.³³ By design, although we might, as imperfect beings, *disagree* about what the epistemically ideal likelihood ratio was for a given item of evidence, there would, in fact, be only one true value.³⁴ Moreover, even if we wanted to identify our disparate estimates of epistemic probability as “noise” and the true epistemic probability as the “signal,” we have little reason to assume that the ideal value always lies at the mean of our flawed estimated values. Accordingly, we must search for some other kind of probability to build our signal/noise analysis upon.

Another candidate is a *frequentist* conception of uncertainty, which is the notion that most of us learned first in school, but this fares no better. To understand frequentist notions of probability, we must first make a few assumptions. There must be a distribution of options from which we are drawing randomly, which is another way to say that we believe we are equally likely to select any particular item in our defined distribution.³⁵ As many scholars have noted, this concept may work well as a means of solving problems in many settings, such as the insurance or gambling industries,³⁶ but it is a poor fit for analyzing the large class of potential evidence items that are not presented in explicitly statistical terms. In theory, we could imagine, for any given item of evidence, taking a survey of all cases that had involved a similar item of evidence, counting the number that ended in guilty and not-guilty findings, and using those percentages as to construct a likelihood ratio. But even if we leave aside for a moment a problem Stein himself has noted, which is that we will rarely get “objective frequentist data” regarding the numerous questions at issue

33. See Walker, *Preponderance*, *supra* note 32, at 1091–92 (distinguishing epistemic probability from subjective probability).

34. Some readers might wish to object that similar evidence might acquire different epistemic probabilities in different cases, but this should be true only due to an ambiguity in our attempt to define what “similarity” means for the purpose of this analysis. Otherwise similar evidence might, for instance, be accorded different weight if the witness had different levels of apparent credibility, but it seems somewhat arbitrary to label such variance a form of noise, rather than a different level of signal for two different types of evidence that we may distinguish between based on finer-grained reference criteria. In other words, the only way to produce variance among epistemic probabilities for a given item of evidence is to hold ourselves ignorant of some epistemically weighty considerations when measuring the average probability as the signal, and then measure the variance produced by taking those extra details into account for the purpose of labeling the evidence’s noisiness.

35. See JOSEPH Y. HALPERN, *REASONING ABOUT UNCERTAINTY* 17–18 (2003).

36. See *id.* at 19.

in litigation,³⁷ we run into insurmountable obstacles. As a basis for a signal/noise analysis, the frequentist conception of uncertainty poses particular problems because it is meaningless to talk about varying frequencies over multiple individual cases. Since the frequentist likelihood ratio for a particular kind of evidence is simply the underlying ratio of guilty and innocent defendants³⁸ in cases involving that evidence, there seems to be no sensible way to conceive of variance in that value.

Nor can we salvage a frequentist conception of the signal by looking for some other form of variance among measures of frequency. It might seem, for a moment, as if we could look at variance among the frequencies we observe in different groupings of cases, but this founders into arbitrariness given the problem of reference classes. Because each case is unique, defining “similarity” with respect to both the cases and evidence items is a judgment call. If the evidence involves a hearsay accusation, should we look at all hearsay accusations? Or just those involving the same crime? Should we restrict our search even further, hunting for similar circumstances giving rise to the statement? Moreover, do we look at all cases in the same jurisdiction? The same country? As we ponder such questions, we will quickly discover that there are a limitless number of reference classes we might examine for any particular item of evidence, and that there are plausible arguments to be made in favor of a large number of candidate classes.³⁹ Here, we face the opposite problem that we encountered above: we can easily conceive competing values for a frequentist likelihood ratio by looking at different reference classes, but it is quite hard to say what basis we would have for choosing a particular one as the optimal “signal.” And even if we try to answer that difficult question, we have no warrant to assume that the mean probability across a broad set of reference classes would match the one drawn from the most appropriate reference class.

But all is not lost. If we instead base Stein’s signal/noise theory on *subjective* likelihood ratios,⁴⁰ it may be possible to make all his

37. See Alex Stein, *Bayesianism Justified*, 1 INT’L J. EVIDENCE & PROOF 339, 341–42 (1997).

38. Or, if we prefer the underlying event definition over the likelihood ratio definition, the ratio of cases involving the fact to cases not involving the fact.

39. See Allen & Pardo, *supra* note 23, at 111–14.

40. The invocation of subjective likelihood ratios may bother some ratios as it seems to raise the ghost of the endless debates about Bayesian theories of evidence. I have no wish to retread that ground in this Response Essay, but I will note two things briefly. First, we might find subjective likelihood ratios to be a useful analytical tool even if we think that they do not provide an accurate account of the ways that jurors ordinarily reason. Second, for the purpose of facilitating a signal-to-noise ratio analysis, one point on which Bayesian models have an advantage is that they offer us numerical descriptions of the strength of individual evidence items. Competing accounts have tended to take a more holistic approach, focusing on the strength and coherence of overall narratives rather than the

assumptions more justifiable. A subjective probability is a measure of a particular person's degree of conviction with regard to a proposition.⁴¹ In the case of subjective likelihood ratios, they would describe the extent to which a particular fact finder viewed a particular evidence item as indicating guilt or innocence.⁴² Common experience suggests that people are often persuaded to different extents by the same item of evidence, so we will have no problem coming up with a range of possible values for subjective likelihood ratios, even if we confine our attention to a small pool such as the members of a particular court or the individuals eligible for jury service in a particular jurisdiction. Thus, it will not be awkward at all to speak of the variance in actual persuasive power as "noise."

Moreover, unlike in the cases of epistemic or frequentist probabilities, there seems to be some useful meaning captured when we consider the signal to be the mean of subjective likelihood ratios over a given population of potential fact finders. Such an estimate might be thought a useful way of describing the persuasive force we should expect the evidence to have in the case before we know enough details about the fact finder to make a more focused and specific estimate. In this framing, the "signal" becomes the average persuasive force of the evidence, as described by the function $S(E)$ defined above, and the "noise" becomes the variance in $S(E)$ judgments that we would expect potential fact finders to make. I do not know if Stein would agree with this interpretation of his theory, but it seems the most sensible way to explain his claims that noise is a measure of the variance in some kind of probability and that the signal is the mean of those varying probabilities.⁴³

impact of each individual evidence item. While I tend to agree that holistic accounts capture more accurately both the subjective experience of fact-finding and some aspects of how judges view the practice, it is hard to use them to produce a numerical score of evidential strength. Readers who doubt, as I do, that judges and juries can or should consider likelihood ratios as a means of analyzing non-statistical evidence items are once again encouraged to think of the ratios as an analytical tool only, used to help us translate the varying judgments that members of the public might have regarding evidential strength into a distribution of values that run from 0 to 1.

41. See Walker, *Preponderance*, *supra* note 32, at 1083–92 (exploring the differences between subjective and epistemic notions of probability in the legal context).

42. See *id.* at 1083.

43. It is also possible that Stein intends to blend the different conceptions of probability in his theory; perhaps noise could be thought of as the variance in subjective probability measures, while signal could be thought of in epistemic or frequentist terms. This might help us make more sense of his normative recommendations, but as I discuss below, it will be hard to square this approach with the assumption that the signal is always found at the mean of the probabilities that give rise to the measure of noise or with the assumption that lower SNR values equate to higher costs attendant to admission.

II. THE SIGNAL-TO-NOISE RATIO AND EVIDENTIAL EFFICIENCY ARE NOT EQUIVALENT CONCEPTS

Having spent some time exploring what, exactly, the concepts of “signal” and “noise” might mean, let us now turn our attention to a particularly important argumentative move that Stein makes. When a particular item of evidence involves a low amount of signal and a high amount of noise, Stein claims it is “inefficient” and will tend to add more expense to fact-finding than can be justified on the basis of its informational value. Therefore, he urges, all evidence with a low SNR ought to be excluded.⁴⁴ Courts can (and do) exclude some items of evidence on efficiency grounds in every judicial system that I am aware of. The equation of low SNR values with inefficiency, however, is less obvious than Stein makes it seem.

To see why the equation of low SNR and inefficiency is doubtful, let us examine closely one of Stein’s examples. He describes a hypothetical case in which a plaintiff is allowed to prove that he suffered severe emotional harm by learning about a tragic accident even though he was not present at the scene.⁴⁵ Because the only evidence of such harm (by assumption) is the plaintiff’s self-serving testimony, Stein claims that there would be wide range of possible “probabilities” that might be assigned to any given case, with an overall signal of 0.5.⁴⁶ Because a jury might find the plaintiff credible and thus find the defendant liable, a trial must be had.⁴⁷ But because the sole item of evidence has a signal that is 0.5, the plaintiff cannot satisfy the preponderance standard and will lose.⁴⁸ If every such trial results in defeat, many social resources will be wasted, and thus, Stein argues, the existing rule of law prohibiting such suits from ever being brought increases overall efficiency.⁴⁹

For discussion purposes, let us go along with Stein for a while in assuming that a signal value of 0.5 requires the plaintiff to lose.⁵⁰ The first problem with this example is that the same result would still follow even if

44. See Stein, *supra* note 3, at 427.

45. See *id.* at 438.

46. See *id.*

47. See *id.*

48. See *id.*

49. See *id.* at 435–39.

50. If readers are willing to go along with the characterizations of signal and noise that I offered in Part I, it would seem that truly equivocal evidence, like the testimony of a possibly self-serving plaintiff on a matter where cross-examination is unlikely to be revealing, should have a much lower signal value. Such testimony has a likelihood ratio very close to 1, and correspondingly it should have a signal value that is very close to 0. This would make it possible to say that the SNR was low even when the noise was relatively small, which seems sensible.

the evidence had a much higher SNR. Assume, for a moment, that the public had learned over time to be wary, but not entirely distrustful, of self-serving testimony, and thus the average $S(E)$ assigned to it ranged, not from 0 to 1, but instead from 0.45 to 0.55, with the same mean of 0.5. Because a 0.55 probability of liability is enough to satisfy the preponderance standard, summary judgment would still be denied, necessitating trials in every case. Likewise, if we assume (as Stein does) the juries would all eventually use the signal value of 0.5 as the basis for their actual verdict, the plaintiffs would still lose each trial, thus consuming a substantial volume of social resources. But by construction, this evidence had a SNR ratio of 10! Clearly there is much more to the idea of evidential efficiency than signal versus noise.

Now notice something else: the majority of the cost savings in Stein's example arose, not from the signal-to-noise relation but from the choice to place the signal value at 0.5. Consider, now, a world in which we are still bad at catching self-serving plaintiffs in their lies, but in which many people also know, based on experience, that truth-telling plaintiffs outnumber lying plaintiffs, in a 3:2 ratio. Now, we might still expect a wide range of judgments with respect to a particular individual, but we might also expect the central tendency of these judgments to be closer to the average frequency of lying, so that we see a range running from 0 to 1 but with a higher overall proportion of potential fact finders finding the evidence to be persuasive, leading to a signal of 0.6. Retaining the assumption that the signal must be the mean of the distribution and its variance is the noise, we might still expect to observe a noise as high as 0.8, leading to an SNR as low as 0.75. But despite this low SNR, two other important things have changed. First, if we continue to assume that jurors will ultimately choose their verdicts based on the signal value, all the plaintiffs will now win instead of lose. And second, although we will be spending more money, it will be to produce liability findings in favor of the party who, we assume, is more likely to be in the right, given that truth-telling plaintiffs do outnumber liars. If, in a world where the above assumptions were true, we were to exclude this testimony on the grounds that it had a low SNR, it would be hard to defend that decision on the grounds of efficiency.

The above discussion shows, I think, that even if SNR and efficiency measures are related, they are correlated rather than equivalent concepts. But I think there are deeper problems in trying to use SNR as a measure of efficiency, which come to light once we start to question some of the assumptions at play in Stein's examples. For instance, Stein frequently seems to assume that, given sufficient resources, judges and juries will typically give evidence only as much force as the signal value would

suggest,⁵¹ but it is far from clear why this would be true under any reasonable conception of what the idea of “signal” is supposed to represent.

If the signal is, as I have characterized it, a function of the average of many subjective probability assessments, then we should not so readily assume that it will always offer reliable predictions of judge and jury decision-making in particular cases. This is most obvious in the case of fact-finding performed by a single official, such as a judge during a bench trial. Consider a judge’s determination regarding the credibility of a witness. In such a circumstance, we might think of the “noise” as the variance among the subjective credibility judgments that differing judges in that jurisdiction might be likely to reach, and the “signal” as the average of those judgments. Why, though, would we expect that each judge on the court would both know what the signal average would be and be willing to defer to it over his own judgments regarding credibility? Certainly nothing in the rules requires him to do so; in fact, trial judges are given great latitude regarding their determinations of witness credibility, even when they rely on fairly noisy indicators such as demeanor evidence.⁵² If we admit that this is a likely possibility, then the low SNR of credibility assessments has a very different outcome. It will not increase the cost of the proceedings by much at all, because judges will quickly default to their own, intuitive assessments of credibility, which will take no longer than assessing evidence that has a higher SNR. But it will, by contrast, increase the expected outcome variance in similarly situated cases. Thus, rather than being uniquely connected to evidential efficiency, we should expect the SNR to also correlate with the reliability and validity of case outcomes.

Jury trials might require a different assessment, however. The more representative the jury is of the larger population, the more likely it will be that the average juror’s judgment regarding the persuasive force of evidence will match the average in the larger population. Since juries nearly always make verdict decisions that line up the initial leanings of a majority of the jurors,⁵³ we might hope that for important items of evidence they will ultimately find the evidence approximately as persuasive as its signal. Here, however, we might also expect that the jury would need to deliberate longer when there is more disagreement among its members

51. See Stein, *supra* note 3, at 435–39 (characterizing the process of fact-finding as one of extracting the reliable signal from the noise at varying levels of cost).

52. See Mark Spottswood, *Live Hearings and Paper Trials*, 38 FLA. ST. U. L. REV. 827, 833–34 n.20 (2011) (collecting citations to rules that require deference to credibility judgments in a broad range of fact-finding contexts).

53. See SAUL M. KASSIN & LAWRENCE S. WRIGHTSMAN, *THE AMERICAN JURY ON TRIAL: PSYCHOLOGICAL PERSPECTIVES* 182–84 (1988) (discussing both observational and experimental studies confirming this tendency).

regarding the value of particular, important items of evidence. In such cases, we could expect at least a modest rise in the cost of litigation without much variance in outcomes. Still, the facts remain that deliberation is a comparatively inexpensive phase of the trial, requiring fewer lawyer work hours than most other stages, and that jurors in most jurisdictions are paid a miserable pittance in exchange for their work. As a result, most cost savings arising from reductions in deliberation time would be modest in scope.

One potential response to this line of thinking is to reject my earlier assumption equating “noise” with the variance of subjective judgments of persuasive power in a population, and “signal” with the mean of those judgments. We might assume, conversely, that “signal” represents not the average judgment in a population, but instead some *objective* indication of evidential value. To make such a description sensible, we would need to determine the signal based on our best available judgment as to the real epistemic value of the evidence as evaluated by maximally competent individual who is ideally rational and resistant to bias. In order to retain Stein’s assumption that fact finders generally accord a weight to evidence that is in line with its signal, we must also assume that real-world fact finders can and do reason optimally, so that in that context, the subjective likelihood ratios match the epistemic ones. Under such assumptions, we might be willing to conclude that noisy evidence requires more cognitive labor from both judges and juries before they arrive at the signal tendency, thus extending the amount of time they must spend deliberating or writing opinions. We might also expect some rises in proof costs when advocates were forced to respond to noisy evidence, such as when it would be necessary to call an otherwise-unnecessary expert witness to clarify matters. In that case, we would expect a modest increase in social cost to arise for each item of noisy evidence introduced during a trial.

There are two problems with such an approach, however. First, recall that we chose the subjective likelihood ratios as the source for our definitions of signal and power precisely to make sense of Stein’s description of the relations between those two terms. If we try and derive both the signal and the noise of evidence from epistemic probabilities, we quickly encounter a roadblock because there is no meaningful way of describing a distribution of objectively ideal assessments of evidential force. To have noise we must have variance, and epistemic probabilities will be single values, not distributions of possible values.

To get around this, we might derive the signal from epistemic likelihood ratios, and the noise from subjective likelihood ratios, but now we will encounter a different problem: we have no general warrant for assuming that the average judgment of persuasive force across a population of people will match its true epistemic value. Psychologists have

documented countless situations in which individual judgments fail to conform to normative rationality, leaving participants with beliefs that are internally or logically inconsistent.⁵⁴ Such failures are not mere artifacts of the laboratory environment. Scholars of naturalized decision-making have similarly identified numerous scenarios in which real-world decision makers, in high stakes situations, fail to reach the right result even though they were experts or had information at hand that indicated that they were making a mistake.⁵⁵ Most chillingly, the same kind of mistakes that are made in some classic laboratory experiments involving undergraduates can also be seen in the retrospective analysis of trials that ended in a wrongful conviction.⁵⁶

When we focus on these last sort of mistakes, we can see clearly the gap that must exist between the signal as a central tendency of typical judgments and the signal as an epistemic ideal. On subjects that come up regularly at trial, common sense intuitions fail to align with normative rationality. People, for instance, regularly place great trust in a confident eyewitness, even though there is ample evidence that confidence and accuracy are poorly correlated and even in circumstances where there is little warrant to believe that the viewing conditions made it likely that the eyewitness formed a clear picture of the events under dispute.⁵⁷ Likewise, people seem to persistently underestimate the possibility that an innocent defendant would confess.⁵⁸ Examples like these are important because they show not just that people can err in either direction relative to a normative inferential baseline, but that in some cases they will tend to err systematically. This will not only happen when decision makers act lazily or fail to think carefully about the subjects under discussion. Rather, as I have argued elsewhere, some kinds of cognitive errors can become magnified through deliberation, even when people have good intentions,

54. See DANIEL KAHNEMAN, THINKING, FAST AND SLOW 3–22 (2011); KEITH E. STANOVICH, RATIONALITY AND THE REFLECTIVE MIND 1–28 (2011).

55. See, e.g., GARY KLEIN, SOURCES OF POWER: HOW PEOPLE MAKE DECISIONS 273–75 (1999); PHILIP E. TETLOCK, EXPERT POLITICAL JUDGMENT: HOW GOOD IS IT? HOW CAN WE KNOW? 128–29, 161–63 (2005).

56. See BRANDON L. GARRETT, CONVICTING THE INNOCENT: WHERE CRIMINAL PROSECUTIONS GO WRONG 5–13 (2011) (tracing the causes of 250 false convictions to causes such as over-reliance on false confessions produced by high-pressure interrogation tactics and misplaced trust in the testimony of confident but mistaken eyewitnesses).

57. See DAN SIMON, IN DOUBT: THE PSYCHOLOGY OF THE CRIMINAL JUSTICE PROCESS 150–57 (2012).

58. *Id.* at 160–62; see also Mark Spottswood, *The Hidden Structure of Fact-Finding*, 64 CASE W. RES. L. REV. 131, 180–93 (2013) (discussing one chilling example in which a jury relied on a confession to convict a defendant of the rape and murder of a child despite the fact that the defendant was excluded by DNA forensics as the source of the semen left in the body of the victim).

given the natural impulse to use our reason to defend, rather than attack, our own intuitive judgments.⁵⁹

Taking these facts into account, it seems very hard to maintain the assumptions that would be necessary to maintain the defense of the SNR-efficiency connection I suggested above. First, the fact that we can observe systematic mistakes in decisions by judges, juries, and other expert decision makers should reduce our confidence that real-world fact finders will usually accord epistemically appropriate weight even to noisy evidence. Instead, we appear to confront an empirical question which is hard to answer: How often, and under what conditions, will they be able to do so? Moreover, the fact that people make these sort of mistakes even in high-stakes scenarios should also make us cautious before assuming that a lower SNR will generally imply higher costs in terms of cognitive and deliberative labor needed to extract the signal from the noise. If people make systematic mistakes, then in some subset of cases, they will expend less effort and give the easy, noisy answer. And once again, we are confronted with an empirical quandary rather than a syllogism: How often, exactly, do fact finders respond to confusing or uncertain evidence with careful analysis, rather than quick, intuitive assessments? I cannot say, and neither, I suspect, can Stein.

To put this theoretic critique in clearer context, let us consider one of Stein's other examples of low-SNR evidence, which he labels as "self-asserting hearsay" testimony.⁶⁰ Stein uses this term for a particular subset of hearsay statements in which the out-of-court statement has no special indicia of reliability and the declarant cannot be called as a witness to be cross-examined for clarification purposes.⁶¹ Stein declares that these statements can have "any probability of being true," and I must agree that it will often be hard to assess the credibility of such statements.⁶² Such statements are, undoubtedly, weak evidence with a low signal value. Moreover, we might also agree, for the sake of discussion, that members of the public will tend to disagree wildly over the likelihood that the statements are true or false, making them also noisy.

Still, even conceding that the SNR ratio of such testimony is low, I am not so sure that we should agree that admitting it is highly inefficient and excluding it correspondingly more efficient. First, it will be the rare case in which a directed verdict motion will be won or lost on the basis of the

59. See Spottswood, *supra* note 58, at 188–93 (describing the cognitive mechanisms that make it easier, in some cases, for jurors to rationalize away problematic evidence items rather than revise judgments of guilt that feel intuitively correct).

60. Stein, *supra* note 3, at 432, 444–45.

61. See *id.* at 444–46.

62. *Id.* at 445.

admissibility of a single piece of self-asserting hearsay. Thus, excluding such testimony will rarely buy us the large efficiency gains discussed in Stein's example of the emotional harm plaintiff mentioned above. Second, admitting such evidence will not add much cost in the form of additional trial time in the typical case, especially when the witness who would testify about it would be called to testify about other matters anyway. Third, the cost of admitting such evidence does not seem to depend very much on its degree of noisiness. Plausibly, even if people disagree over whether they trust or distrust such evidence, no one trusts or distrusts it without reservation to the extent that they would stake their lives on it. Rather, ordinary common sense suggests that all such statements must be viewed with caution. Ergo, even if the variance of opinions regarding such evidence is greater than its probative force, it is very unlikely that a typical jury would spend much time debating the question. Rather, they would acknowledge, just as they would in everyday life, that such weak sources of information are not worth trusting very much to support conviction or exoneration.⁶³ We might say, in terms of Stein's theory, that such evidence has a signal of 0.1 (very close to a likelihood ratio of 1, suggesting minimal probative value) and a noise of 0.1, suggesting that, at best, jurors will expect it to be a little bit likely to be either true or untrue, but in no event very powerful. Thus, we get a SNR of 1, which is low, but it is also hard to envision jurors spending much additional deliberation time on this evidence. Adding it all up, it would seem that although the evidence is weak and the noise is high, the costs are also minimal.

Finally, though, and most importantly, Stein neglects a different kind of cost which is produced by the need to operate a complicated mechanism like the modern hearsay rule. The lines between self-asserting and non-self-asserting hearsay are not always clear, and in fact, practicing lawyers spend great amounts of time and effort arguing for or against the exclusion of various items of hearsay evidence.⁶⁴ If, as Stein suggests,⁶⁵ this is to be justified in terms of cost savings, we must also tally up, on the other side of the equation, all the many costs that such a complicated rule imposes. These include the cost of training law students, year after year, how to understand one of the most complicated legal rules ever created; the cost of testing their knowledge on the bar examination; the time that lawyers must spend arguing over the admissibility of hearsay items when trying to settle cases or negotiate pre-trial orders; the additional motion practice and

63. See Michael L. Seigel, *Rationalizing Hearsay: A Proposal for a Best Evidence Hearsay Rule*, 72 B.U. L. REV. 893, 910–11 (1992).

64. See Ronald J. Allen, Commentary, *The Evolution of the Hearsay Rule to a Rule of Admission*, 76 MINN. L. REV. 797, 800 (1992).

65. Stein, *supra* note 3, at 444–45.

judicial work required to resolve the disputes they cannot negotiate away; the time and expense incurred whenever judges must give, and juries must work to follow, complicated instructions to use a particular out-of-court statement for one purpose but not for another. Adding all these costs up, I doubt very much that they could be outweighed by the small cost of the extra trial time incurred when a witness repeats the occasional bit of self-asserting hearsay or the modest additional deliberation time that might be required before the jury decides to give that testimony little weight. All of which is to say, even if the SNR value of an evidence item tells us *something* about the costs it might impose on the system, there are surely a number of other facts that are as much, if not more, important.

III. WHY THE TURN TOWARDS EFFICIENCY ANALYSIS IS ESSENTIAL

Readers who have come this far might doubt, frankly, my statement in the introduction that I find this to be one of the most important evidence papers I have read in the last few years. I have, after all, disagreed both with the way in which Stein characterizes his signal-to-noise ratio as applied to evidence items, as well as with his central claim that low SNR evidence should generally be excluded because it is inefficient. Despite those concerns, however, Stein has done something truly important with this work, which is to encourage the rest of the evidence academy to redirect their attention from the questions that have typically occupied our attention towards the question of evidential efficiency.⁶⁶ Despite my reservations about the specific theory of efficiency that Stein has offered, I agree wholeheartedly that evidential efficiency is a topic both important and neglected.

Evidence scholarship all too often focuses on a standard set of questions. Many writers, among whom I would include myself, devote extensive attention to analyzing how particular evidence rules might either increase or decrease the accuracy of trial outcomes in individual cases.⁶⁷ Other writers focus more on preventing particular kinds of outcome errors (such as wrongful convictions) or on assessing the fit between current doctrines and the sources of authority that give rise to them.⁶⁸ Finally, another vein of scholarship attempts to assess the subtler effects that evidence rules might have, such as by causing distress to crime victims or influencing the incentives people have to engage in different kinds of

66. *Id.* at 424.

67. See Koehler, *supra* note 22, at 880–84.

68. See SIMON, *supra* note 57, at 206–22.

primary conduct.⁶⁹ In this landscape, a systematic assessment of the relationship between different evidentiary rules and the cost of trying cases is, quite frankly, a breath of fresh air. I do not say this because I have any special love of accountancy or tallying up sums. Rather, a turn to efficiency analysis is crucial because the costs and the outcome accuracy of a litigation system go hand-in-hand.

Stein describes one serious problem with allowing cost to go uncontained in a large system of justice, which is the multiplicative impact of a small amount of waste across a large volume of cases.⁷⁰ We might add to that kind of concern a second one, which is the effect that high costs may have on the frequency with which parties take cases to trial.⁷¹ It is commonly understood by litigation scholars that raising trial costs increases the incentives of parties to resolve their cases via settlements and plea bargains.⁷² Such behavior is no doubt rational for many parties who obtain a net surplus by avoiding trial costs and who can split some of that surplus between them when bargaining. But that analysis only makes sense when we are holding the costs of trial static. If we can find ways of making trials faster and cheaper without significantly undermining their accuracy, many parties might then find that the costs of taking cases to trial no longer outweigh their benefits. Indeed, such considerations no doubt help to explain the rising popularity of arbitration among parties who find themselves frequently in court. By making it easier to take cases to trial, we can be assured that those cases that end in settlements and plea bargains are more likely to be handled in that way because the parties prefer a negotiated outcome rather than because that is the only kind of outcome they can afford.

Moreover, when trials are hard to obtain, many other kinds of harm can follow. Law's ability to command compliance depends on a number of factors, including the availability and salience of public information regarding the consequences of law-breaking, as well as the system's overall legitimacy in the eyes of the public.⁷³ Excessively high-settlement and plea-bargaining rates undermine the first objective. Many civil settlement bargains are made and kept in secret, and as a result, when the number of settlements goes up, the number of public determinations regarding the line

69. See Tom Lininger, *Prosecuting Batterers After Crawford*, 91 VA. L. REV. 747, 768–71 (2005).

70. See Stein, *supra* note 3, at 433–35.

71. Cf. Marc Galanter, *The Vanishing Trial: An Examination of Trials and Related Matters in Federal and State Courts*, 1 J. EMPIRICAL LEGAL STUD. 459, 465, 494, 527 (2004) (tracing the decline of civil and criminal trials in federal and state courts over time).

72. See, e.g., George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1, 3–6 (1984).

73. See, e.g., TOM R. TYLER, *WHY PEOPLE OBEY THE LAW* 64–68 (1990).

between legal conduct and liable conduct falls.⁷⁴ Plea bargains must be made publicly, but nevertheless, such proceedings lack the salience and drama of a criminal trial that ends in a verdict. Finally, both surveys and observational studies have revealed a strong connection between perceptions of courts' legitimacy, on the one hand, and the public's willingness to follow the law's commands, on the other.⁷⁵ The trial ritual is the justice system's public face and the setting in which we go farthest to demonstrate our respect for the dignity of both the victim and the accused. As such rituals shrink from public visibility, we may find that the public's respect and obedience to the law diminish as well. Conversely, to the extent we can find reforms in procedure and evidence law that makes trials both affordable and fair, we may increase not only the satisfaction of parties in litigation, but overall compliance with, and respect for, the law.

CONCLUSION

Professor Stein's article is a welcome breath of fresh air, in that it invites us as evidence scholars to lift our gaze from the particular and consider the impact of admitting items of evidence at a systemic level. We should all be grateful for this invitation, because it is all too easy to get caught up in the complex analysis required to determine what the accuracy effects of a piece of evidence might be within the confines of individual cases. In such analysis, it is standard to take certain shortcuts, such as assuming that the case will be taken to trial either way and that the parties and the justice system will not be affected in any broader way by the decision. We make such assumptions at our peril.

I have, in this Response, offered some focused criticisms of the specific way that Stein implemented the ideas of "signal" and "noise" as descriptors of evidence items and of his linkage between those metrics and the broader idea of evidential efficiency. It should not be surprising that there is room for improvement on these technical matters because Stein's project was massive and he took it very far, very quickly. I make these criticisms in the hope that he and other scholars of proof will continue to examine the cost effectiveness of evidential rules—a course of study that may greatly enrich our shared field of inquiry.

74. See Janet Cooper Alexander, *Do the Merits Matter? A Study of Settlements in Securities Class Actions*, 43 STAN. L. REV. 497, 567 (1991).

75. See TYLER, *supra* note 73, at 62–64.