*Rebecca K. Helm, Valerie P. Hans & Valerie F. Reyna**

Legal cases often require jurors to use numerical information. They may need to evaluate the meaning of specific numbers, such as the probability of match between a suspect and a DNA sample, or they may need to arrive at a sound numerical judgment, such as a money damage award. Thus, it is important to know how jurors understand numerical information, and what steps can be taken to increase juror comprehension and appropriate application of numerical evidence. In this Article, we examine two types of juror decisions involving numbers––decisions in which jurors must convert numbers into meaning (such as by understanding numerical evidence in order to determine guilt or liability), and decisions in which jurors must convert meaning into numbers (such as by understanding qualitative evidence and converting this into a numerical damage award amount). In each of these areas we analyze legal cases and research to examine areas in which dealing with numbers leads to sound or sub-optimal decision making in jurors. We then examine psychological theory and research on numerical decision making to understand how informed, fair, and consistent juror decision making about numbers can be promoted. We conclude that what is often most important is juror understanding of the meaning of numbers in context rather than technically precise numerical ability, supporting the role of the lay jury. We also suggest how to improve juror understanding, so that jury decisions better reflect considered community judgment.

* Rebecca K. Helm, LLM, Cornell Law School, 2011, is a Lecturer in Law at the University of Exeter; Valerie P. Hans, Ph.D., is a Professor of Law at Cornell Law School, Cornell University; Valerie F. Reyna., Ph.D., is a Professor in the Department of Human Development, and is associated with the Human Neuroscience Institute and the Center for Behavioral Economics and Decision Research, Cornell University. Preparation of this Article was supported in part by the Martha E. Foulk Fellowship awarded to Rebecca K. Helm, by National Science Foundation award SES-1536238: "Quantitative Judgments in Law: Studies of Damage Award Decision Making" to Valerie P. Hans and Valerie F. Reyna, by a grant from Cornell University's Institute for Social Sciences to Valerie P. Hans and Valerie F. Reyna, and by National Institute of Health (National Institute of Nursing Research) award RO1NR014368-01 to Valerie F. Reyna. We thank David Kaye for his insightful comments on the manuscript. Correspondence concerning this Article should be addressed to Valerie P. Hans, Cornell Law School, Ithaca, NY 14853. E-mail: valerie.hans@cornell.edu.

INTRODUCTION

Jurors are generally sound decision makers in both civil and criminal trials. Findings from empirical studies show that the strength of evidence presented at trial is a major determinant of jury verdicts, and that civil jury damage awards are strongly correlated with the degree of injury in a case.[1] However, despite the overall strengths of juries, some cases place significant demands on citizen jurors. These cases include those in which jurors are required to engage in numerical decision making, either through considering numerical evidence or through being asked to reach a numerical decision such as a percentage of liability or a damage award amount. The information presented to jurors in these cases is becoming increasingly complex as new scientific techniques and technological innovations develop and as their presentation in the nation's courtrooms becomes more detailed and complicated.[2] Litigators

---

[1] Shari Seidman Diamond, *Beyond Fantasy and Nightmare: A Portrait of the Jury*, 54 BUFF. L. REV. 717 (2006); Valerie P. Hans & Neil Vidmar, *The Verdict on Juries,* 91 JUDICATURE 226, 226 (2007); HARRY KALVEN, JR. & HANS ZEISEL, THE AMERICAN JURY (1966); Richard O. Lempert, *The American Jury System: A Synthetic Overview*, 90 CHI.-KENT L. REV. 825 (2015); SUJA THOMAS, THE MISSING AMERICAN JURY: RESTORING THE FUNDAMENTAL CONSTITUTIONAL ROLE OF CRIMINAL, CIVIL, AND GRAND JURIES (2016); NEIL VIDMAR & VALERIE P. HANS, AMERICAN JURIES: THE VERDICT (2007).

[2] *See* Sanjeev Bajwa, *Apple v. Samsung: Is it Time to Change our Patent Trial System?* 27 GLOBAL BUS. & DEV. L.J. 77 (2014); EXECUTIVE OFFICE OF THE PRESIDENT, PRESIDENT'S COUNCIL OF ADVISORS ON SCIENCE AND TECHNOLOGY, REPORT TO THE PRESIDENT: FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF FEATURE-COMPARISON METHODS (Sept. 2016), https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf.

are increasingly presenting detailed numerical information to jurors, moving away from the previous practice of offering simple conclusory statements by experts.[3]

The complexity of numbers that jurors must now understand and use is illustrated by the recent patent litigation involving two high-tech giants, Apple and Samsung. In this case, a civil jury was asked to determine whether Samsung had infringed on numerous Apple patents, and if so, the sum of damages that Samsung should pay Apple for this infringement.[4] Jurors were required to delve deeply into complex financial documents and expert opinions to determine if patents were infringed, and if so, what they should be worth. This included determining the precise profits for twenty-eight devices and determining the percentage of profits owed for the infringement.[5] This task was further complicated by the fact that numbers varied greatly depending on whether they were generated by Apple's or Samsung's experts.[6] This was clearly a complex task, as the jury tackled a twenty-page verdict form requiring hundreds of judgments.[7] Following the trial, the judge ruled that some damages needed to be recalculated due to jury error.[8]

Concerns about the jury's ability to understand, critically evaluate, and apply complex evidence, including numerical evidence, have led to arguments that "blue-ribbon" juries should be used in some cases.[9] Blue-ribbon juries are juries selected for their special qualities such as advanced education or special training.[10] Such juries are controversial since they are drawn from unrepresentative samples of the population.[11] In ad-

3 William C. Thompson & Eryn J. Newman, *Lay Understanding of Forensic Statistics: Evaluation of Random Match Probabilities, Likelihood Ratios, and Verbal Equivalents,* 39 LAW & HUM. BEHAV. 332, 332 (2015).

4 William F. Lee, WilmerHale, Presentation to Cornell Law School, The Smartphone Wars: The Apple-Samsung Trial (Feb. 20, 2017); Greg Sandoval, *How Qualified is the Apple-Samsung Jury? We Found Out,* CNET (Aug. 24, 2012, 2:52 PM), http://news.cnet.com/8301-13579_3-57499944-37/how-qualified-is-the-apple-samsung-jury-we-found-out.

5 Nilay Patel, *Apple vs. Samsung: Inside a Jury's Nightmare,* THE VERGE (Aug. 23, 2012, 10:31 AM), http://www.theverge.com/2012/8/23/3260463/apple-samsung-jury-verdict-form-nightmare.

6 *Id.*

7 Apple Inc. v. Samsung Electronics Ltd., Samsung Electronics America Inc., Samsung Telecommunications America, LLC., No. 11-CV-01846-LHK, verdict form (N.D. Cal. August 21 2012), http://assets.sbnation.com/assets/1307288/1890_finalverdictform.pdf.

8 Lee, *supra* note 4; Josh Lowensohn, *Apple v. Samsung: Judge Orders New Trial on Some Damages, Cuts Award by $450M,* CNET (Mar. 1, 2013, 12:50 PM), https://www.cnet.com/news/apple-v-samsung-judge-orders-new-trial-on-some-damages-cuts-award-by-450m.

9 Jordan M. Halle, *Avoiding Those Wearing Propeller Hats: The Use of Blue Ribbon Juries in Complex Patent Litigation,* 43 U. BALT. L. REV. 435, 436 (2014); JAMES OLDHAM, TRIAL BY JURY: THE SEVENTH AMENDMENT AND ANGLO-AMERICAN SPECIAL JURIES (2006).

10 *Id.*

11 *See* OLDHAM, *supra* note 9; VIDMAR & HANS, AMERICAN JURIES, *supra* note 1, at 66–69 (describing historical changes in opinions about special and blue-ribbon juries).

dition, studies suggest that special juries of highly educated individuals may not be a panacea for the challenges created by complex trials.[12] Although such jurors may be better educated and better able to understand numbers and perform mathematical calculations, they may still lack understanding of nuances and the meaning of numbers in a particular type of case.[13] Research has shown that even highly educated laypersons have a relatively poor understanding of probabilities, risks, and other chance-related concepts.[14] Blue-ribbon jurors are likely to be about as susceptible as the rest of us to common cognitive errors and biases in numerical decision making.[15] To minimize errors when jurors deal with numerical information, an understanding of how people process numerical information and how this can lead to cognitive errors and bias is essential.

This Article begins by summarizing what we currently know about how juries understand and use numbers in legal decision making, first in the context of moving from numbers to a meaningful qualitative judgment, and second in the context of translating a qualitative judgment into a numerical decision. Then, it discusses theory-based research on numerical cognition that can provide insight into how accurate and consistent use of numbers can be facilitated. Finally, it draws on this theory-based research to make suggestions for assistance that should be provided to juries dealing with numbers in the trial setting. We analyze the positive and negative effects of selecting juries with high levels of numeracy. We conclude that numeracy is not a panacea in helping jurors when dealing with numbers. In the majority of cases, what is most important is juror understanding of the meaning of numbers rather than technical and precise numerical ability. Importantly, technical numerical ability is unlikely to reduce many common juror errors and could potentially increase certain types of bias. This leads us to provide four new suggestions to help jurors dealing with numbers: (1) targeted training; (2) use of visual aids; (3) the provision of information to give context to numbers; and (4) active jury trial reforms.

---

12 For discussion of studies showing problems encountered by highly educated and numerate people, *see* Section IV.A *infra*.

13 Bajwa, *supra* note 2, at 90.

14 *See* Isaac M. Lipkis, Greg Samsa, & Barbara K. Rimer, *General Performance on a Numeracy Scale Among Highly Educated Samples,* 21 MED. DECISION MAKING 37, 39–41 (2001).

15 *See* Jeffrey J. Rachlinski, Andrew J. Wistrich & Chris Guthrie, *Can Judges Make Reliable Numeric Judgments? Distorted Damages and Skewed Sentences*, 90 IND. L.J. 695, 706–10 (2015) (showing that judges are influenced by anchoring effects); Valerie F. Reyna, Wendy L. Nelson, Paul K. Han & Nathan F. Dieckmann, *How Numeracy Influences Risk Comprehension and Medical Decision Making,* 135 PSYCHOL. BULL. 943, 946 (2009).

I. EXTRACTING MEANING FROM NUMBERS

Jurors are often asked to evaluate evidence involving probability and statistics. One common moment is when a forensic scientist is attempting to explain the strength of a piece of forensic evidence presented at a jury trial.[16] In the past, such evidence was presented in a categorical fashion (e.g. two fingerprints were made by the same finger, they were not made by the same finger, or the test was inconclusive).[17] In 1997, the Federal Bureau of Investigation announced that due to technological advances, "their experts would be permitted to testify that DNA from blood, semen, or other biological evidence at a crime scene came from a specific person."[18] This practice was criticized as reducing the potential donor pool to one source; the FBI claim was characterized as "not possible," and "not logically attainable."[19] However, it could be reasonably argued that reporting a source attribution within that donor pool when the probability of another person having the same DNA profile in the relevant population is nearly zero indeed uses the logic of probability to make a claim that is very probably true (given the assumptions in the analysis). The bottom line is that some forms of criminalistics evidence are presented quantitatively, and we can expect increasing efforts to do so with more identification methods being developed in the future, with ensuing confusions and debates about probability judgments.

Recently, forensic scientists have been encouraged to characterize the value of the evidence using statistics such as probabilities.[20] For example, in the case of DNA evidence, when a suspect's DNA matches a DNA sample from a crime scene, the probative value of that DNA evidence may be conveyed by the probability of finding a match in someone who is randomly sampled from a reference group, known as the random match probability ("RMP").[21] A low RMP means that the match between a suspect and a recovered sample is unlikely to have occurred by chance; instead, it is more likely that the defendant was truly the source of the recovered sample. In contrast, a higher random match probability means a match between a suspect and a recovered sample is more likely to be coincidental, that is, to have occurred by chance. RMPs are frequently

---

16 Thompson & Newman, *supra* note 3.

17 Simon A. Cole, *Individualization is Dead, Long Live Individualization! Reforms of Reporting Practices for Fingerprint Analysis in the United States,* 13 LAW, PROBABILITY & RISK 117, 121–22 (2014).

18 Constance Holden, *DNA Fingerprinting Comes of Age,* 278 SCIENCE 1407, 1407 (1997).

19 Cole, *supra* note 17, at 122.

20 Thompson & Newman, *supra* note 3; Mike Redmayne, Paul Roberts, Colin Aitken & Graham Jackson, *Forensic Science Evidence in Question,* 5 CRIM. L. REV. 347, 355–56 (2011).

21 Cole, *supra* note 17, at 132–33.

used by the prosecution in this way in cases to convey the significance of a match.[22]

Providing jurors with such statistics may allow them to appropriately weigh evidence. Research suggests that jurors are often influenced appropriately and as intended by statistical information.[23] For example, research on RMPs shows that juror evaluations of evidence do vary appropriately when the RMP is varied (more weight is given to the evidence when the RMP is lower and less weight is given to the evidence when the RMP is higher).[24] However, jurors do face challenges when dealing with numbers in this context. Psychological theory and existing research suggest that jurors are at risk of being subject to cognitive errors and inconsistencies when interpreting statistical evidence, that jurors may frequently be unable to critique complex statistics given by experts effectively, and that jurors' existing viewpoints may influence their interpretations of statistical evidence.[25]

A. *Interpreting Statistical Evidence*

Evidence suggests that the way in which numerical evidence is framed and presented can influence how jurors think about the evidence and how probative they consider it to be.[26] This means that jurors may be convinced that the same statistical evidence is impressive or insufficient.[27] Like the rest of us, jurors are susceptible to specific cognitive biases when evaluating statistical evidence, which can result in inconsistent interpretations of identical evidence and errors in interpreting statistical evidence.[28]

In the context of DNA RMPs, Jonathan Koehler has identified two ways that changing the presentation of statistics can change juror inter-

---

[22] *See* United States v. Shea, 957 F. Supp. 331, 345–46 (D.N.H. 1997); *People v. Xiong*, 215 Cal. App. 4th 1259, 1269–70 (2013); *United States v. Morrow*, 374 F. Supp. 2d 51, 54 (D.D.C. 2005).

[23] David L. Faigman & A. J. Baglioni Jr., *Bayes' Theorem in the Trial Process: Instructing Jurors on the Value of Statistical Evidence,* 12 LAW & HUM. BEHAV. 1, 16 (1988); Brian C. Smith, Steven D. Penrod, Amy L. Otto & Roger C. Park, *Jurors' Use of Probabilistic Evidence,* 20 LAW & HUM. BEHAV. 49, 78 (1996); William C. Thompson, Suzanne O. Kaasa & Tiamoyo Peterson, *Do Jurors Give Appropriate Weight to Forensic Identification Evidence?,* 10 J. EMPIRICAL LEGAL STUD. 359, 359 (2013).

[24] *Id.*

[25] *Cf* .Dan M. Kahan, Ellen Peters, Maggie Wittlin, Paul Slovic, Lisa Larrimore Ouellette, Donald Braman & Gregory Mandel, *The Polarizing Impact of Science Literacy and Numeracy on Perceived Climate Change Risks*, 2 NATURE CLIMATE CHANGE 732, 732–33 (2012).

[26] Jonathan J. Koehler, *The Psychology of Numbers in the Courtroom: How to Make DNA-Match Statistics Seem Impressive or Insufficient,* 74 S. CAL. L. REV. 1275, 1277 (2001).

[27] *Id.* at 1282–84.

[28] COLIN AITKEN & FRANCO TARONI, STATISTICS AND THE EVALUATION OF EVIDENCE FOR FORENSIC SCIENTISTS 81–82 (2004) (describing some common errors).

pretation of evidence—through changing the target of the statistics and through changing the frame.[29] When considering RMPs, the "target" of the DNA match can either be the suspect or a larger reference population.[30]

Imagine a case in which a DNA match has been found between a suspect and material recovered from a crime scene. Consider the following two pieces of evidence:

a) The chance that *the suspect* would match the DNA if he were not the source is 1 in 1,000.
b) The DNA of 1 in every 1,000 *people in the city in which the crime occurred* would match the DNA.

When the target is the subject, as in (a), it seems less likely that a match occurred by chance, but when a larger reference population is the subject, as in (b), it seems more likely that a match occurred by chance. Although these pieces of evidence are equivalent, Koehler has shown that jurors are likely to perceive DNA evidence as less incriminating when the RMP is presented as it is in (b).[31] Jurors who are encouraged to think about a larger reference group (e.g. people in the city in which the crime occurred), rather than the suspect, are likely to be less persuaded by DNA evidence suggesting that the suspect is guilty.[32] This, in turn, influences judgments of guilt.[33]

The way that information is framed, in particular whether the evidence is framed as a frequency or a probability, has also been shown to influence how people think about statistical evidence.[34] Jurors are more likely to be convinced by DNA evidence when RMPs are presented as probabilities (e.g. a .001 chance of a random match) rather than frequencies (e.g. a 1 in 1,000 chance of a random match).[35] This finding is also true for numbers more generally—people tend to overestimate frequencies compared to probabilities.[36]

Research has examined approaches other than RMPs to convey the importance of a DNA match.[37] One study using mock juror participants compared three ways of reporting the reliability of evidence indicating a match between the defendant and evidence from a crime scene (focusing

[29] Koehler, *supra* note 26.
[30] *Id.*
[31] Jonathan J. Koehler, *When Are People Persuaded by DNA Match Statistics?* 25 LAW & HUM. BEHAV. 493, 497–98 (2001).
[32] *Id.*
[33] *Id.* at 502.
[34] *Id.* at 497
[35] *Id.* at 502
[36] Valerie P. Hans & Valerie F. Reyna, *To Dollars from Sense: Qualitative to Quantitative Translation in Jury Damage Awards,* 8 J. EMPIRICAL LEGAL STUD. 120, 124 (2011).
[37] *See* Thompson & Newman, *supra* note 3.

on DNA evidence and shoeprint evidence)––using RMPs, likelihood ratios, and verbal equivalents.[38] Using likelihood ratios involves considering two mutually exclusive hypotheses––H1: that the items have the same source, and H2: that the items do not have the same source.[39] The likelihood of the observed results occurring under each of the two hypotheses is then calculated and the ratio of these two likelihoods is reported.[40] This can result in the conclusion that the probability of obtaining this evidence if the items have the same source is *x* times higher than obtaining this evidence if the items do not have the same source.[41] So, the ratio gives the likelihood of a DNA match if the defendant was the true source of the DNA evidence, compared to the likelihood of a DNA match if the defendant was not the true source of the DNA evidence. A likelihood ratio greater than one means it is more likely that the defendant was the true source of the DNA evidence, while a likelihood ratio of less than one means it is more likely that the defendant was not the true source. (Note that a likelihood *ratio* greater than one provides no assurance that the probability that the defendant was the true source is greater than .50). Using verbal equivalents involves computing a likelihood ratio and then reporting the results in words rather than numbers, according to a graduated scale.[42]

A 2013 study found that, as predicted, participants were influenced by the way in which the forensic evidence was presented; RMP statements helped jurors to distinguish best between strong and moderate forensic evidence.[43] For example, participants viewing very strong shoeprint evidence were more influenced by that evidence when shown RMPs than when shown likelihood ratios or verbal equivalents.[44] Participants viewing shoeprint evidence were only sensitive to the strength of the evidence (very strong or moderate) when shown RMPs, and not when shown likelihood ratios or verbal equivalents.[45]

In addition to inconsistencies when interpreting evidence, mock jurors have been shown to make cognitive errors. One example is the source probability error. This is the tendency to equate the chance that a suspect would match a DNA trace if he was not the source with the

38 Thompson & Newman*,* *supra* note 3.

39 *Id.* at 333.

40 *Id.*

41 *Id.*; Geoffrey Stewart Morrison, *Measuring the Validity and Reliability of Forensic Likelihood-Ratio Systems,* 51 SCI. & JUST. 91, 97–98 (2011).

42 Kristy A. Martire, Richard I. Kemp, Ian Watkins, Malindi A. Sayle & Ben R. Newell, *The Expression and Interpretation of Uncertain Forensic Science Evidence: Verbal Equivalence, Evidence Strength, and the Weak Evidence Effect,* 37 LAW & HUM. BEHAV. 197, 197 (2013).

43 *Id.* at 203–04.

44 *Id.* at 205.

45 *Id.*

probability that the suspect is not the source of the trace.[46] For example, one might assume that:

a) The chance that the suspect would match the DNA if he were not the source is 1 in 1,000, means that;
b) The chance that the suspect is not the source of the trace is 1 in 1,000.

These seem on the surface to be similar, but some careful reflection will show why they are not equivalent. Although the DNA evidence can include the defendant in a relatively small group of potential sources, it cannot distinguish the defendant from other individuals in that group and so cannot indicate the probability that the defendant, and not another group member, is the source.[47] Although this is incorrect, people mistakenly infer the probability that matching items have (or do not have) a common source from the RMP. If the chance that the suspect would match the DNA was 1 in 1,000 (as in the example above), there would be around 320,000 people in the USA who would be a match (as the population of the USA is approximately 320 million), so the DNA alone cannot indicate the probability that the defendant was the source of the DNA.

A similar cognitive error, known as the prosecutor's fallacy, occurs when people equate the RMP with the probability that a defendant is innocent given a match.[48] This is statistically incorrect because it equates the probability of being innocent given a match with the probability of a match given innocence (two different values). In addition to being statistically incorrect, there are other sources of potential error too. The report of a match could be the result of a false positive lab error or the planting of incriminating evidence by the police. Defendants have raised the possibility of jurors committing the prosecutor's fallacy to argue against the admissibility of random match probability statistics, but this argument is often not successful.[49]

Evidence also suggests that jurors may have difficulty combining two different reasons why DNA evidence may not be conclusive, for example, the chance that a DNA match was coincidental and the chance that the DNA match was due to a laboratory error.[50] A study by Jason Schklar and Shari Seidman Diamond found that the majority of mock jurors used an improper method to combine such probabilities (such as

---

46 Jonathan J. Koehler, *Error and Exaggeration in the Presentation of DNA Evidence at Trial,* 34 JURIMETRICS J. 21, 27 (1993).

47 Thompson & Newman, *supra* note 3, at 335.

48 Thompson et al., *supra* note 23, at 362.

49 *See* United States v. Shea, 957 F. Supp. 331, 345–46 (D.N.H. 1997).

50 Jason Schklar & Shari Seidman Diamond, *Juror Reactions to DNA Evidence: Errors and Expectancies,* 23 LAW & HUM. BEHAV. 159, 178–82 (1999).

averaging across the two probabilities), leading them to underestimate the combined risk of incorrectly identifying a defendant.[51]

Clearly, it is important that jurors do not fall victim to these fallacies when evaluating evidence, which can interfere with jurors' abilities to comprehend the significance of numerical evidence and to draw appropriate conclusions from it. It is also important that jurors receive appropriate information to be able to effectively critique the evidence presented to them.

B. Difficulty Critiquing Statistics Given by Experts

A second issue that arises when jurors are considering statistical evidence is that jurors face difficulties in evaluating claims made by experts in cases in which complex statistical issues bear on the validity of the experts' conclusions. One area in which this is important is in picking a "reference population" when calculating probabilities. This is a defined population which is used to calculate the probability of an event occurring. In the case of DNA evidence and RMPs, this would be the population chosen to determine the probability of a match occurring coincidentally. For example, if it is reported that there is a 1 in 1,000 chance that a Caucasian male would match the DNA, then the reference group is Caucasian males. There are complex issues involved in choosing a reference group, for example, in the case of DNA evidence, what should be done when there is no information about the ethnic group of the source of a DNA trace? And what should be done when there is DNA from more than one source at a crime scene?[52]

Without additional help, jurors are not necessarily equipped to evaluate whether an appropriate reference group has been picked, and how sound the resulting conclusions are likely to be, which can lead to acceptance of the expert's conclusions about misleading statistical evidence. For example, in one UK case, Sally Clark, a mother with two infant children who had died suddenly, stood trial on charges that she had murdered them.[53] It was initially thought that the babies had died from sudden infant death syndrome. A distinguished pediatrician, Dr. Meadow, estimated that the probability of two children dying of sudden infant death syndrome was around one in 73 million.[54] This was widely interpreted as meaning that the chance the mother was innocent was one in 73 million (a variation of the prosecutor's fallacy described above).[55]

---

51 *Id.*

52 *See* Koehler, *supra* note 46, at 26; D. H. Kaye, *Logical Relevance: Problems with the Reference Population and DNA Mixtures in People v. Pizarro,* 3 LAW, PROBABILITY & RISK 211, 211 (2004).

53 R v. Clark [2003] EWCA Crim. 1020.

54 *Id.* at para. 96.

55 *Id.* at para. 105–08 (discussing this in the context of the appeal decision).

After Dr. Meadow gave evidence that the chance of two sudden infant death syndrome deaths in a family like hers was one in 73 million, he went on to state:[56]

> [I]t's the chance of backing that long odds outsider at the Grand National, you know; let's say it's an 80 to 1 chance, you back the winner last year, then the next year there's another horse at 80 to 1 and it is still 80 to 1 and you back it again and it wins. Now here we're in a situation that, you know, to get those odds of 73 million you've got to back that 1 in 80 chance four years running . . . the chance of it happening four years running we all know is extraordinarily unlikely. So it's the same with these deaths. You have to say two unlikely events have happened and together it's very, very, very unlikely.[57]

This statistic is incorrect for several reasons. The pediatrician simply took the chance of such a death occurring once (1/8,500), and multiplied it by itself to get the probability of the event occurring twice. He mistakenly used the probability of such a death occurring in the general population to calculate the probabilities of the first and the second deaths of the children, incorrectly assuming that the events (the two deaths) were independent. In fact, the probability for the second event should have been calculated by looking at the probability of such a death in families who have already had one case of sudden infant death. These families could justifiably be in a high risk group as there could well be environmental and/or genetic risk factors in the family.[58] Despite the fact that this statistic is incorrect, the evidence was given without objection from the defense,[59] and was not later identified as erroneous by the appeal court (despite the appeal court recognizing that important qualifications were not given when presenting the evidence).[60]

Sally Clark was convicted by the jury. Her conviction was later overturned on appeal as it transpired that one of her sons had colonization of staphylococcus aureus bacteria, indicating that he had died from natural causes, but this evidence had not been disclosed to the defense.[61]

---

56 *Id.* at para. 96.

57 *Id.* at para. 99.

58 *See* Letter from the President of the Royal Statistical Society to the Lord Chancellor regarding the use of statistical evidence in court cases, (Jan. 23, 2002), http://www.rss.org.uk/Images/PDF/influencing-change/rss-use-statistical-evidence-court-cases-2002.pdf.

59 Lisa Claydon, *Law, Neuroscience, and Criminal Culpability*, *in* LAW AND NEUROSCIENCE 141, 164 (Michael Freeman ed., 2011).

60 *See* Letter from the President of the Royal Statistical Society to the Lord Chancellor regarding the use of statistical evidence in court cases, *supra* note 58.

61 R v. Clark [2003] EWCA Crim. 1020.

At the appeal, the court also referenced the statistical evidence that had been given in the first trial, stating that important qualifications were not referred to by Dr. Meadow in this evidence to the jury and "thus it was the headline figures of 1 in 73 million that would be uppermost in the jury's minds with the evidence equated to the chances of backing four 80 to 1 winners of the Grand National in successive years."[62] This statistical error, the inability of the jury to detect the error when evaluating the evidence, and the failure of the defense attorney and the court to recognize and highlight this error, likely led to the erroneous conviction.

Testifying experts like Dr. Meadow do make errors when dealing with statistical evidence. In a 2003 article, Jonathan Koehler lists thirteen cases from the late 1980s and early 1990s in which experts, the courts, or both committed the source probability error described above.[63] For example, in *United States v. Jakobetz*, it was stated that "The FBI . . . calculated that there was one chance in 300 million that the DNA from the semen sample could have come from someone in the Caucasian population other than Jakobetz"[64] and in *People v. Lindsey* it was stated that "[a] genetic epidemiologist . . . testified . . . that the odds of someone besides the defendant having the banding pattern appearing in the known sample and in the forensic sample was one in 340 billion."[65]

Being able to critically evaluate statistical evidence given by experts is important for jurors. Arguably, it is even more important that lawyers and judges are able to critically evaluate statistical evidence given by experts. Lawyers are tasked with presenting and testing the evidence through examination and cross-examination, and judges play a key gatekeeping role.

C. Being Influenced by Existing Viewpoints

Finally, research suggests that all of us, including jurors, are influenced by existing viewpoints when interpreting legal and scientific evidence.[66] The phenomenon of motivated cognition occurs when decision makers have a preference regarding the outcome of an evaluative task, and are therefore more likely to arrive at that outcome by engaging in inadvertently biased processes for accessing, constructing, and evaluating beliefs.[67] This phenomenon has been shown to apply to jurors evalu-

---

62 *Id.* at para. 102.

63 Koehler, *supra* note 46, at 28.

64 United States v. Jakobetz, 955 F. 2d 786, 789 (2d Cir. 1991).

65 People v. Lindsey, 868 P.2d 1085, 1087 (Colo. App. 1993).

66 Avani Mehta Sood, *Motivated Cognition in Legal Judgments––An Analytic Review*, 9 ANN. REV. L. SOC. SCI. 307, 308–09 (2013).

67 *Id.*; *see* Kahan et al., *supra* note 25.

ating expert presentation of scientific evidence including RMPs.[68] Specifically, research has shown that mock jurors with high levels of education are more likely to be influenced by political and social viewpoints when evaluating DNA evidence.[69] This could be important in how jurors evaluate statistics.

In sum, interpreting and critiquing statistical evidence in order to evaluate the strength of evidence presented in a jury trial can be challenging. The intrinsic demands of the task are further complicated by well-known heuristics, biases, and the phenomenon of motivated cognition.

II. EXTRACTING NUMBERS FROM MEANING

A far less explored area in which jurors have to deal with numbers occurs when jurors have to generate numbers from their assessment of qualitative evidence. This occurs most often when jurors make compensatory and punitive damage awards. It also occurs when they determine the proportion of liability in contributory and comparative negligence cases, and when they, rarely, are asked to undertake criminal sentencing. This conversion of qualitative to quantitative judgments is central to many legal decisions, especially in the civil context.[70] Researchers have found some reassuring evidence that jurors are generally able to rank order the severity of injuries and the seriousness of crimes.[71] For example, compensatory damage awards are correlated with the severity of the plaintiff's injury, punitive damage awards are correlated with the size of compensatory awards, and criminals who commit serious offenses generally spend more time in prison than those who commit less heinous crimes.[72]

Despite this, jurors typically make assessments about individual cases in isolation, without looking at comparable cases, so ensuring that similar cases are treated alike is left to judges and the appeals process. In addition, these cases naturally involve a large amount of ambiguity because of limited guidance from the court and the difficulty of assigning

---

68 Rebecca K. Helm & James P. Dunlea, *Motivated Cognition and Juror Interpretation of Scientific Evidence: Applying Cultural Cognition to Interpretation of Forensic Testimony,* 120 PENN ST. L. REV. 1, 2–3 (2016).

69 *Id.*

70 *See* Valerie P. Hans, Jeffrey J. Rachlinski & Emily G. Owens, *Editors' Introduction to Judgment by the Numbers: Converting Qualitative to Quantitative Judgments in Law,* 8 J. EMPIRICAL LEGAL STUD. 1, 3 (2011).

71 *See* Theodore Eisenberg, Jeffrey J. Rachlinski & Martin T. Wells, *Reconciling Experimental Incoherence with Real-World Coherence in Punitive Damages,* 54 STAN. L. REV. 1239, 1240–41 (2002).

72 *Id.*

value to factors that are often intangible.[73] This has led to criticisms of the variability, unpredictability, and bias in juror decisions.[74] These tasks would likely be challenging for any decision maker,[75] and jurors recognize the difficulty of these tasks.[76] Thus, it is important to identify what the challenges are for jurors, and how these challenges can be ameliorated.

A. *Inconsistency and Unpredictability of Juror Awards*

It has been argued that the translation of moral judgments into dollars and years is not grounded in principle, and so it produces large differences among people.[77] Numerical awards made by jurors are often highly variable, even when jurors agree on qualitative aspects of a case.[78] One experimental study using jury-eligible respondents examined judgments of the outrageousness of a defendant's actions and also judgments of punitive dollar awards made by individuals and by synthetic juries.[79] This study showed that there was substantial consensus among jurors on the outrageousness of a defendant's actions and the appropriate severity of punishment, but that judgments of dollar awards were much more variable.[80] Notably, even those who agreed on the outrageousness of the defendant's actions and the appropriate severity of punishment differed in their dollar award amounts.[81] This convergence on qualitative judgments and variability in dollar award recommendations has also been shown in other work.[82]

Researchers have identified multiple sources of ambiguity involved in reaching numerical judgments in such cases including limited gui-

---

73 Valerie F. Reyna, Valerie P. Hans, Jonathan C. Corbin, Ryan Yeh, Kelvin Lin & Caisa Royer, *The Gist of Juries: Testing a Model of Damage Award Decision Making,* 21 PSYCHOL. PUB. POL'Y & L. 280, 280–81 (2015).

74 Eisenberg, Rachlinski & Wells, *supra* note 72.

75 *See* KENNETH R. FEINBERG, WHAT IS LIFE WORTH?: THE INSIDE STORY OF THE 9/11 FUND AND ITS EFFORT TO COMPENSATE THE VICTIMS OF SEPTEMBER 11TH (2006) (describing the difficulty of assigning monetary value to the loss of life).

76 *See* Shari Seidman Diamond, *What Jurors Think: Expectations and Reactions of Citizens Who Serve as Jurors*, *in* VERDICT: ASSESSING THE CIVIL JURY SYSTEM, 282 (Robert E. Litan ed., 1993).

77 Cass R. Sunstein, Daniel Kahneman, David Schkade & Ilana Ritov, *Predictably Incoherent Judgments*, 54 STAN. L. REV. 1153, 1155–56 (2002).

78 Daniel Kahneman, David Schkade & Cass R. Sunstein, *Shared Outrage and Erratic Awards: The Psychology of Punitive Damages,* 16 J. RISK & UNCERTAINTY 49, 49–50 (1998).

79 *Id.*

80 *Id.*

81 *Id.* at 50.

82 Roselle L. Wissler, Allen J. Hart & Michael J. Saks, *Decisionmaking About General Damages: A Comparison of Jurors, Judges, and Lawyers*, 98 MICH. L. REV. 751, 757–58 (1999).

dance, uncertainty of projections, and intangible losses.[83] Another ambiguity can result when jurors do not understand the meaning of the numbers given in context (for example that $100,000 is a small amount of punitive damages to a company like the fast food corporation McDonald's). This means jurors can understand a case, but end up giving an arbitrary damage award amount.[84] This ambiguity likely contributes to the substantial variance observed in both judge and juror damage awards.[85] This is exacerbated by the fact that jurors, unlike judges, consider cases in isolation and do not have comparable cases to refer to.[86] In fact, in many cases lawyers are actually barred from referring to awards in other cases.[87] This can lead to inconsistency in numerical awards across cases. For example, someone who has endured less pain and suffering might receive a higher damage award than someone who has endured more pain and suffering.[88]

Juries face a particularly ambiguous situation in determining monetary damages in personal injury cases. For example, in tort cases, juries are tasked with putting plaintiffs in the position that they would have been in had the tort not occurred. This involves assessing and quantifying both economic and non-economic losses, some of which are intangible. What, for example, is the numerical worth of becoming incontinent, or of losing function in one lung? Dollar awards are an unbounded scale; although the minimum award is set at zero, there is no maximum award amount. Judgments made on unbounded scales tend to be more variable than judgments made on bounded scales with pre-set minimum and maximum levels.[89]

The unpredictability built into damage awards creates opportunities for unequal treatment of similar cases, and makes it difficult for insurance agents, lawyers, and others who worry about their inability to predict jury awards. In turn, that makes it difficult to assess in advance whether to take a case to trial or settle to avoid disastrous awards.[90] This may also harm litigants, whether the harm is to plaintiffs who receive very low awards that do not cover the costs of their injuries, or the harm is to defendants who may be ordered to pay substantial judgments be-

---

[83] Shari Seidman Diamond, Mary R. Rose, Beth Murphy & John Meixner, *Damage Anchors on Real Juries,* 8 J. EMPIRICAL LEGAL STUD. 148, 148–49 (2011).

[84] Reyna et al., *supra* note 73.

[85] *Id.*; Hans & Reyna, *supra* note 36.

[86] Sunstein et al., *supra* note 77, at 1156.

[87] *Id.*

[88] *Id.* at 1155.

[89] Kahneman et al., *supra* note 78.

[90] Jane Goodman-Delahunty, Maria Hartwig, Pär Anders Granhag & Elizabeth F. Loftus, *Insightful or Wishful: Lawyers' Ability to Predict Case Outcomes,* 16 PSYCHOL., PUB. POL'Y, & L. 133, 133–34 (2010).

yond their ability to pay.[91] On the other hand, some commentators have pointed out that the comparative unpredictability of trial decisions increases the power of the courts and the jury to deter wrongful and harmful behavior.[92]

B. Biasing Influences

Biasing factors can detrimentally affect numerical award amounts given by jurors, even when they are irrelevant to the case. The well-documented anchoring heuristic offers a case in point.[93] Anchors are numbers that can readily bias numerical judgments by being used as a starting point that people fail to adjust from sufficiently.[94] Thus, subjects who see a high number (a high anchor) before giving a damage award amount have been shown to give larger awards than those who see a low number (a low anchor) before giving a damage award amount.[95] This is also the case in sentencing decisions made by judges.[96] One study showed that providing criminal trial judges with unrealistically high sentencing options led them to subsequently give longer sentences to defendants in criminal cases, even when they were told the suggestion of the original option was a mistake.[97]

Virtually all anchoring research has been conducted using meaningless anchor numbers.[98] People have been asked to consider whether an appropriate estimate is above or below one's telephone number or roulette wheel number, for example.[99] Recent research has extended the study of anchors to look at "meaningful" anchors.[100] These are anchors that are meaningful and relevant in determining what a damage award amount should be. For example, information on the average annual income may convey useful information about the likely value of a particu-

---

[91] Valerie P. Hans, Rebecca K. Helm & Valerie F. Reyna, From Meaning to Money: Translating Injury into Dollars (unpublished manuscript) (on file with authors).

[92] Stephan Landsman, *Introduction*, 60 DEPAUL L. REV. 271, 271 (2011) ("[L]awyers abhor uncertainty . . . [but] this denigrates the value of uncertainty, something that may be critical to the deterrence of misconduct . . .").

[93] *See* DANIEL KAHNEMAN, THINKING, FAST AND SLOW, 124–26 (2011).

[94] Gretchen B. Chapman & Eric J. Johnson, *Anchoring, Activation, and the Construction of Values,* 79 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 115, 115–16 (1999).

[95] *See* Diamond, et al., *supra* note 83, at 152–53; Verlin B. Hinsz & Kristin E. Indahl, *Assimilation to Anchors for Damage Awards in a Mock Civil Trial*, 25 J. APPLIED SOC. PSYCHOL. 991, 992–94 (1995); Mollie W. Marti & Roselle L. Wissler, *Be Careful What You Ask For: The Effect of Anchors on Personal Injury Damage Awards,* 6 J. EXPERIMENTAL PSYCHOL. APPLIED 91, 92 (2000); Reyna et al., *supra* note 73, at 288.

[96] Birte Englich & Thomas Mussweiler, *Sentencing Under Uncertainty: Anchoring Effects in the Courtroom,* 31 J. APPLIED SOC. PSYCHOL. 1535, 1540–41 (2001).

[97] *Id.* at 1546–47.

[98] For a summary of research on anchoring, *see* KAHNEMAN, *supra* note 93, at 117.

[99] Chapman & Johnson, *supra* note 94, at 116.

[100] Reyna et al., *supra* note 73, at 283–84.

lar dollar award amount to a plaintiff. Importantly, meaningful anchors have been shown to influence mock jurors more than meaningless anchors.[101] This suggests that when jurors do have reference points that should be relevant to their decisions, these will be more influential than meaningless numerical amounts.

Juror numerical awards have also been shown to be influenced by juror characteristics, such as financial status and juror attitudes towards civil litigation. For example, higher income jurors tend to recommend larger dollar awards,[102] and jurors who are more critical of civil litigation tend to give lower awards.[103] The influence of numerical anchors and individual juror characteristics on quantitative judgments made by jurors suggests that such judgments from one case to the next could vary, dependent upon the use of anchors and the makeup of the jury. That could lead to inequitable outcomes.

Another area in which jurors are required to make numerical judgments is when they are required to consider contributory or comparative negligence of the plaintiff in a civil lawsuit. Tort law provides that a defendant can raise a plaintiff's own failure to act reasonably as a defense in a negligence action. Damages awarded to the plaintiff can be reduced or even eliminated based on this negligence.[104] In these cases, jurors must determine the proportion of the blame attributable to the plaintiff's own negligence.[105] In the United States, jurisdictions either employ a pure comparative negligence approach, whereby a plaintiff's recovery is reduced in accordance with the plaintiff's share of responsibility regardless of what that proportion is, or a modified comparative negligence approach, where a plaintiff is barred from recovery when the plaintiff's negligence reaches 50% or 51%.[106]

Research has shown that the extent to which the jury finds the plaintiff to have been negligent is influenced by the law they are being asked to apply.[107] In one study examining awards from 2001–2005, researchers found that in pure comparative negligence jurisdictions, juries found that the plaintiff's responsibility exceeded 50% in 22% of cases, while in modified comparative negligence jurisdictions, the plaintiff's responsibility was only found to exceed 50% in 7.5% of cases.[108] When making

---

101 *Id.* at 285–86.

102 Wissler et al., *supra* note 82, at 783.

103 Valerie P. Hans, *What's It Worth? Jury Damage Awards as Community Judgments,* 55 WM. & MARY L. REV. 935, 952–53 (2014).

104 JENNIFER K. ROBBENNOLT & VALERIE P. HANS, THE PSYCHOLOGY OF TORT LAW 150 (2016).

105 *Id.* at 150–52.

106 *Id.*

107 Eli K. Best & John J. Donohue III, *Jury Nullification in Modified Comparative Negligence Regimes,* 79 U. CHI. L. REV. 945, 945–46 (2012).

108 ROBBENNOLT & HANS, *supra* note 104, at 154.

their determinations, jurors are clearly influenced by the recovery they believe would be just, rather than strictly determining the numerical proportion of plaintiff liability, which can shift their numerical judgments. It is also likely that in modified comparative negligence states when numbers around 50% are specifically mentioned as cut-offs, that these numbers will serve as anchors, influencing determinations of plaintiff responsibility.[109]

To minimize errors and biases when jurors deal with numerical information, an understanding of how people process numerical information and how this can lead to cognitive errors and bias is essential. As in Section I, which illustrated the challenges of moving from numbers to a meaningful qualitative judgment, Section II summarized problems confronting decision makers when they must translate their qualitative judgments into numerical decisions. We now turn to relevant psychological theory and potential remedies.

III. THEORY AND RESEARCH ON NUMERICAL DECISION MAKING

Psychological theory and research provides insight into people's ability to understand and use numerical information in decision making. First, and unsurprisingly, numeracy has been shown to improve the ability to deal with numbers in many contexts. For example, people will be unable to make accurate judgments if they do not know what a decimal point means, or what the "/" means in a fraction.[110] This can lead to inaccurate conclusions, such as 1/3 + 1/2 = 1/5.

Ellen Peters and her colleagues conducted four studies comparing individuals who were high or low in numeracy on decisions that required understanding of probability and statistics.[111] The highly numerate individuals were more likely than those with low numeracy to understand and employ appropriate numerical principles, which made the highly numerate people less susceptible to framing effects.[112] Those lower in numeracy were affected more by irrelevant or affective factors.[113] In one study, for example, participants read about a psychiatric hospital patient and were asked to judge how much risk it would pose to release the patient.[114] Half of them read the "frequency" version ("Of every 100 patients similar to Mr. Jones, 10 are estimated to commit an act of vio-

---

109 *Id.* at 157–58.

110 *See* Valerie F. Reyna & Charles J. Brainerd, *The Importance of Mathematics in Health and Human Judgment: Numeracy, Risk Communication and Medical Decision Making,* 17 LEARNING & INDIVIDUAL DIFFERENCES 147, 147–48 (2007).

111 Ellen Peters, Daniel Västfjäll, Paul Slovic, C.K. Mertz, Ketti Mazzocco & Stephan Dickert, *Numeracy and Decision Making*, 17 PSYCHOL. SCI. 407 (2006).

112 *Id.* at 408.

113 *Id.* at 412–13.

114 *Id.* at 409.

lence to others during the first several months after discharge"); the other read the "percentage" format (Of every 100 patients similar to Mr. Jones, "10% are estimated" to commit an act of violence to others during the first several months after discharge).[115] The actual risk (10 of 100, or 10%) is equivalent. Compared to those lower in numeracy, those with strong number skills were better able to identify the equivalency of the probability and frequency frames in the presentation of potential risk, and to use the insight appropriately in decision making.[116]

However, it is clear that numeracy is not a panacea.[117] More recent research suggests that what is most important is understanding the meaning of numbers, rather than just details and arithmetic.[118] As we will discuss below, this research provides insight into how numerical information should be presented to jurors, and suggests that systematic biases could actually be introduced by selecting only highly numerate jurors.

We focus here on fuzzy-trace theory, a psychological theory of memory and decision making that predicts and explains many findings in the literature on numeracy and decision-making, particularly in the medical decision-making context.[119] Fuzzy-trace theory distinguishes between verbatim and gist representations of information. Verbatim representations capture the surface form of information (for example exact words or numbers), while gist representations capture its meaning or interpretation (based on context, and other factors known to affect meaning).[120] So, for example, take an individual who is considering bungee jumping. A 10% chance of rain may be considered a low chance in the context of deciding whether to perform a bungee jump, but a 10% chance of death would likely be considered a high chance. Similarly, a $100,000 punitive damage award may be considered a huge amount for a private individual defendant but only a small amount for a corporate defendant. Research has shown that people encode both verbatim and gist representations of information separately, and that each typically forms the basis of a different kind of reasoning––one focused on precise details (verbatim-based processing) and the other on understanding global meaning (gist-based processing).[121]

There are multiple stages to encoding information and using it to make a decision, and therefore there are various stages at which errors can be introduced into decision making. The information must be heard, and encoded as a representation, the information must then be retrieved

---

115 *Id.*
116 *Id.*
117 *See* Reyna et al., *supra* note 15, at 943.
118 *See* Reyna et al., *supra* note 73, at 281.
119 Reyna et al., *supra* note 15, at 965–66.
120 *Id.* at 965.
121 *Id.*

appropriately, and finally the information must be processed and integrated, along with other information, as part of the decision-making process. When an individual is making a decision, errors can therefore occur in multiple ways.

First, errors can occur due to a lack of knowledge or a failure to encode information in the appropriate way. For example, a juror might believe a statistic given to them by an expert because they do not have sufficient knowledge themselves to critique the information that the expert is giving, as in the Sally Clark case. In addition, a juror who did not understand what the "/" in a fraction meant would not be able to accurately draw conclusions from evidence given in the form of fractions. Inconsistencies in damage award amounts can also be caused by a lack of background knowledge giving jurors little on which to base their decisions. Errors that occur due to a lack of knowledge can be cured by providing jurors with the necessary knowledge that they need to understand and interpret information. Errors can also occur due to a failure to retrieve appropriate knowledge in context. These errors can be cured by cueing the knowledge relevant to properly deal with numbers.[122]

However, errors can, and often do, occur due to processing interference rather than a lack of knowledge or failure to retrieve knowledge. These errors occur despite jurors having the relevant knowledge to avoid biases, due to confusion when information is being processed. A set of common errors occurring due to processing interference are class-inclusion errors. Class-inclusion errors can occur when people deal with overlapping sets and subsets. For example, if 3/100 people are a DNA match, we have an overlapping set and subset since the subset of 3 in the numerator is also included in the overall set of 100. These overlapping numbers are confusing and can interfere with processing, which in turn can lead to biases such as focusing on the numerator in a fraction and neglecting the denominator. For example, people show a tendency to judge a low probability event as more likely when presented as a large numbered ratio (a 10/100 chance of an event occurring) than when it is presented as an equivalent smaller numbered ratio (for example, 1/10), despite knowing that the two are numerically equivalent.[123] Processing errors are often simple bookkeeping errors that can be cured by presenting information clearly to avoid confusion. Although greater numeracy helps to the extent that understanding a number is necessary to have a basic understand-

---

[122] *See* C. J. Brainerd & V. F. Reyna, *Autosuggestibility in Memory Development,* 28 COGNITIVE PSYCHOL. 65, 96–97 (1995); *see* V. F. Reyna & C. J. Brainerd, *Fuzzy-Trace Theory: An Interim Synthesis,* 7 LEARNING & INDIVIDUAL DIFFERENCES 1, 35–44 (1995).

[123] *See* Valerie F. Reyna & Charles J. Brainerd, *The Origins of Probability Judgment: A Review of Data and Theories, in* SUBJECTIVE PROBABILITY 255 (George Wright & Peter Ayton eds., 1994).

ing of what it means, these errors are committed even by numerate individuals.[124]

Many of the errors that have been shown to occur when jurors deal with numbers are processing errors.[125] To illustrate, let's return to the case of RMPs. As described above, jurors are more likely to be convinced by DNA evidence when RMPs are presented as probabilities (e.g. a .001 chance of a random match) rather than frequencies (e.g. a 1 in 1,000 chance of a random match). This is likely to be because they neglect the denominator in the frequency and focus on .001 and 1. As .001 is much lower than 1, the chance of a RMP seems smaller.

Class-inclusion errors are also responsible for more complex errors, such as the prosecutor's fallacy (when people equate the RMP with the probability that the defendant is innocent), and the difficulty in dealing with two potential sources of error. These errors can be compared to errors in assessing risk of a disease, errors that have been shown to occur due to overlapping classes. Take the following example:

> The pretest probability of a disease is 10% and a diagnostic test has 80% sensitivity (80% of people with the disease test positive) and 80% specificity (80% of people without the disease test negative). Is the probability that a specific individual who has a positive test result has the disease closer to 30% or 70%?

Ignoring the denominator here (which is far smaller for the group having the disease than the group who do not have the disease due to the pretest probability of 10%) leads to confusing the probability of a positive result given the disease, with the probability of the disease given a positive result. This means that most people (even medical experts) say that the probability is closer to 70%, when the correct answer is closer to 30%.[126] This error can be corrected by using 2x2 tables or Venn diagrams (diagrams using overlapping circles to show different classes and their areas of overlap) to separate classes from one another. For example, the above problem can be illustrated in Table 1, the 2x2 table below:

---

[124] Brainerd & Reyna, *supra* note 122, at 19.

[125] Valerie F. Reyna, *How People Make Decisions That Involve Risk: A Dual-Process Approach,* 13 CURRENT DIRECTIONS IN PSYCHOL. SCI. 60 (2004).

[126] *Id.* at 64–65.

TABLE 1. TWO-BY-TWO TABLE CLEARLY SEPARATING CLASSES TO PROMOTE ACCURATE REASONING WHEN INTERPRETING OUTCOMES OF BREAST CANCER SCREENING

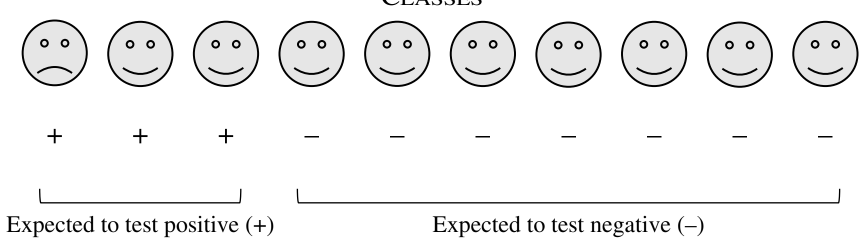
	Chance of positive test	Chance of negative test	Totals
Breast cancer	8	2	10
No breast cancer	18	72	90
Totals	26	74	100

Table 1 makes it clear that if a test result is positive, the chance of having breast cancer is 8/26 (around 30%) and the chance of not having breast cancer is 18/26 (around 70%). This information can also be communicated using simple visual aids. Figure 1 illustrates the following scenario:

> Ten percent of people in a population have a disease. A test is available to detect the disease. This test has 100% sensitivity (100% of people with the disease test positive) and 80% specificity (80% of people without the disease test negative).

In Figure 1, a sad face indicates a person with the disease, and a happy face indicates a person without the disease. As ten percent of the people in the population have the disease, we start off with one out of ten faces being sad, and nine out of ten faces being happy.

FIGURE 1. VISUAL AID TO REDUCE CONFUSION FROM OVERLAPPING CLASSES



Note. Sad face represents a person with the disease; happy face represents a person without the disease. Plus sign indicates a positive test result; minus sign indicates a negative test result.

Because only 80% of people without the disease test negative, 20% of people without the disease test positive. This means of nine people without the disease, we would expect around seven to test negative and two to test positive (as we can see in Figure 1). Therefore, as Figure 1 makes clear, in this scenario, even if someone receives a positive test result, it is more likely than not that they do not have the disease.

In order to reduce errors, it is important to identify how they arise. Errors resulting from a lack of knowledge can often be reduced by providing background knowledge and training. While background information can also be important in preventing processing errors, how information is presented has also been shown to be vitally important in reducing these errors.[127] Understanding how people process information can provide insight into how these errors arise, and how they can be reduced.

When people are making decisions, they have a tendency to rely on gist (simple and meaning-based) rather than verbatim (detailed and surface-level) representations.[128] According to fuzzy-trace theory, reliance on gist has important advantages in decision making as it involves insightful processing based on understanding a stimulus, rather than just precise and analytical calculation. It is likely true that in some cases, for example, damage award cases in which awards for potential future damage are calculated by the probability of that damage occurring multiplied by the value of that damage, or cases in which percentages of profits must be calculated, precise numerical knowledge will be important. However, in many other cases, what will be more important is an understanding of the meaning of numbers in context. Is a $100,000 award for pain and suffering high or low, given the context? Likewise, is a 10% chance of having a random DNA match high or low? Put simply, what is most important in many cases is understanding the meaning of numbers in a given context, rather than precise numerical values. Reliance on meaning (rather than precise and detailed numbers) has been shown to reduce processing errors, including the class-inclusion errors described above.[129] Therefore, what is likely to be very important in interventions is ensuring jurors can understand what each number means in the context in which it is presented.

This is also the case in damage award decision making. A model of juror damage award decision making based on these and other insights from fuzzy-trace theory has been proposed.[130] The Hans-Reyna model posits that jurors first make simple gist judgments, specifically categorical (damages are warranted or not) and ordinal (the damages deserved are low, medium, or high). Jurors then search for a dollar award amount that fits the gist judgments. It is this conversion of gist judgments to precise dollar amounts that is most influenced by heuristics, as jurors rely on anchors or symbolic numbers from everyday life that have meaning to them personally as low or high numbers to assign a numerical value to the gist of the case.[131] This is consistent with research showing that ju-

---

127 Reyna et al., *supra* note 15, at 958–59.

128 *Id.* at 966.

129 *Id.*

130 Reyna et al., *supra* note 73, at 282–84.

131 *Id.,* at 291.

rors tend to converge on evaluations of qualitative aspects of a case, but not damage award amounts.[132] Importantly, the Hans-Reyna model also predicts that meaningful anchor amounts will be more influential than meaningless anchor amounts. This is because meaningful numbers provide greater insight into whether a particular amount is low or high. This prediction has been supported by experimental research.[133]

We now consider ways to help jurors dealing with numbers, based on this research on numerical cognition.

IV. WHAT CAN HELP JURORS DEALING WITH NUMBERS?

A. *Numeracy*

As noted above, unsurprisingly, strong numerical skills can help when making decisions involving numbers. This is likely to be the case where errors are prone to occur due to a lack of knowledge. For example, knowledge of multiplication will help jurors calculating damages for potential future harm (typically calculated by multiplying the risk of the future harm by its magnitude), and knowledge of appropriate reference groups can provide assistance to jurors critiquing expert evidence. Assembling juries that are highly numerate to consider numerical evidence or to deliver numerical verdicts may be considered one way to improve how jurors cope with numbers. As we noted earlier, this can be done by assembling "blue-ribbon" juries drawn from highly educated or trained subgroups of the population in complex cases. More highly educated juries can also be formed as a consequence of jury selection, if lawyers base their peremptory challenges on education and training. In the *Apple v. Samsung* trial, for instance, potential jurors were extensively questioned to assess their likely understanding of the complex evidence they would hear at trial.[134]

Juror numeracy can serve as a partial protective factor for decision making, which is encouraging news. However, fuzzy-trace theory and associated research suggests that this is not a wholly sufficient strategy to solve all problems that arise when jurors deal with numbers. First, research suggests that numeracy does not make a difference in many types of legal decisions involving numbers. This is because the errors and biases shown by jurors when dealing with numbers are usually associated with the processing of information, rather than a lack of knowledge.[135] Research shows that even highly educated and numerate individuals are

132 Kahneman et al., *supra* note 79, at 49.

133 Reyna et al., *supra* note 73, at 285–86.

134 Bajwa, *supra* note 2, at 91.

135 Reyna et al., *supra* note 125.

not immune to making processing errors.[136] When it comes to damage awards, the decisions of more numerate jurors do not necessarily differ from those of less numerate jurors, even when numbers are involved. For example, experimental research to date has found no influence of numeracy on damage award amounts.[137] Recent research has also supported the idea that more numerate jurors may find certain numerical judgments more difficult than less numerate jurors (specifically in the converting meaning into numbers cases), by showing that in a mock personal injury case where a damage award amount for pain and suffering needed to be given, jurors with higher numeracy reported that it was more difficult to decide on a damage award.[138]

In addition, there is new, albeit limited, research evidence that in some highly charged scientific areas, selecting more numerate jurors could paradoxically increase biases in juror decision making. Empirical research on scientific evidence about climate change suggests that as individuals become more scientifically literate, they can also become more biased when evaluating polarizing scientific evidence.[139] This research suggests that more knowledgeable individuals could be more prone to motivated cognition, perhaps due to an increased ability to interpret evidence to fit a pre-existing viewpoint.[140]

Furthermore, research suggests that it is vital for jurors to understand the meaning of numbers in context as well as their quantitative value.[141] Some people may have an ability to perform numerical calculations, but not to understand the meaning of numbers in context.[142] Selecting jurors based on numerical ability could lead to juries with good technical skills, but with a lack of insight beyond the calculation of numbers. Thinking literally by focusing on accurate technical details may interfere with assessing gist and meaning, for example assessing the gist of the pain and suffering of another person in order to convert this into a damage award.[143] Anticipating an injured person's career trajectory and valuing the worth of an individual's life or pain and suffering, as juries

---

136 *See* Gretchen B. Chapman & Jingjing Liu, *Numeracy, Frequency, and Bayesian Reasoning,* 4 JUDGMENT & DECISION MAKING 34 (2009); Reyna, *supra* note 126; Valerie F. Reyna & Mary B. Adam, *Fuzzy-Trace Theory, Risk Communication, and Product Labeling in Sexually Transmitted Diseases,* 23 RISK ANALYSIS 325 (2003).

137 *See* Brian H. Bornstein, *The Impact of Different Types of Expert Scientific Testimony on Mock Jurors' Liability Verdicts,* 10 PSYCHOL. CRIME & L. 429 (2004); Hans et al., *supra* note 91.

138 Hans et al., *supra* note 91.

139 Helm & Dunlea, *supra* note 69; Kahan et al., *supra* note 25, at 732.

140 *Id.*

141 Valerie F. Reyna & Charles J. Brainerd, *Dual Processes in Decision Making and Developmental Neuroscience: A Fuzzy-Trace Model,* 31 DEVELOPMENTAL REV. 180 (2011).

142 *Id.*

143 *See id.* (giving further information and explaining the relationship between autistic traits and meaning-based processing).

are required to do, require insights and judgments that go beyond the mere calculation of numbers. Reflecting on the judgments he had to make as special master of the Congressionally-mandated 9/11 fund to allocate money to the many victims and their families who had suffered losses in the terrorist attacks, Kenneth R. Feinberg said, "It was a job that called for the wisdom of Solomon, the technical skill of H&R Block, and the insight of a mystic with a crystal ball. I was supposed to peer into that crystal ball, consider the ebbs and flows that made up a stranger's life, and translate all of this into dollars and cents."[144] A jury drawn from a representative cross-section of the community is better able to reflect the broad range of local perspectives, and the jury's decision in turn will better reflect the community's norms and values.[145]

A final but very significant problem with blue-ribbon juries is that they do not fully represent the population at large. The evaluation of trial evidence and the assessment of credibility of litigants and witnesses draw importantly on the decision makers' knowledge, past experiences, and perspectives. Selecting jurors based on education alone would remove significant sources of insight from the jury that decides the case.

B. *Training and Warnings*

Another approach is to attempt to educate jurors during a trial, and when necessary, to alert them to potential biases and errors.[146] It is important to ensure that jurors are capable of understanding numbers that are involved in a case, that jurors are walked carefully through all numerical calculations, and that jurors are provided with appropriate information with which to critique expert opinions based on numbers. This is particularly important in preventing errors based on knowledge and retrieval. For example, it may allow jurors to spot errors in expert testimony or to more accurately calculate profits from a financial spreadsheet.

There is also evidence that processing errors can be prevented by providing cues to assist jurors in avoiding those errors.[147] For example, by reminding jurors that a 1/10 risk is the same as a 10/100 risk, inconsistencies in the ascription of meaning to these fractions can be reduced, as jurors will recognize that treating them differently is illogical.

---

[144] FEINBERG, *supra* note 75, at 87.

[145] Hans, *supra* note 103, at 963–64.

[146] Anne W. Reed, *The ABCs of Tutoring Your Jury,* 42 TRIAL TECHNIQUES 20 (2007) (proposing that we should attempt to train prospective jurors in advance of trial and eliminate those who fail to improve); *See* Jonathan J. Koehler, *Train Our Jurors* (Faculty Working Papers, Paper 141, 2006), http://scholarlycommons.law.northwestern.edu/facultyworkingpapers/141.

[147] *See* Reyna, *supra* note 126.

Judges have advocated the educational approach to dealing with numerical challenges in trial evidence.[148] For example, in one case considering the risk of the prosecutor's fallacy in jurors dealing with random match probabilities, the United States District Court in New Hampshire stated "if a trial judge concludes that jurors could be confused by statistical evidence, the judge can deliver carefully crafted instructions to insure that the evidence is properly understood."[149]

However, it is important to note that general warnings are unlikely to be a panacea for all of the issues that surround dealing with numbers. In order to be effective, warnings and training have to be carefully designed based on the errors that jurors may be susceptible to in a particular case. In the Sally Clark case discussed above, the judge in the first case did give a warning regarding the statistical evidence, he stated:

> I should, I think, members of the jury just sound a word of caution about the statistics. However compelling you may find them to be, we do not convict people in these courts on statistics. It would be terrible day if that were so. If there is one SIDS (sudden infant death syndrome) death in a family, it does not mean that there cannot be another one in the same family.[150]

Although this warning did alert jurors that a defendant should not be convicted based on statistics, it was relatively vague and so would be unlikely to give the jurors the knowledge they would need to critique the statistics effectively and to reject the incorrect figure offered by the prosecution's expert.

Statistical training may be valuable for jurors. But it may be even more essential for judges and attorneys so that they can recognize when calling an expert in statistics may be important, and when jurors are likely to need guidance. We've illustrated the potential biasing effects of framing the presentation of statistical information in particular ways. Thus, training and warnings should be accompanied by presenting numerical evidence to jurors in a way that helps them in processing that information and drawing appropriate conclusions from it. Below we consider some ways that the meaning of numbers can be conveyed to jurors, and how this can help to avoid processing errors and inconsistencies.

C. Providing Meaning Through Visual Aids

As noted above, informed decisions depend on being given accurate information, and also through truly comprehending the meaning of that

---

148 *See* United States v. Shea, 957 F. Supp. 331, 345 (D.N.H. 1997).
149 *Id.*
150 R v. Clark [2003] EWCA Crim. 1020, at para. 104.

information. Many errors that occur when jurors deal with numbers are due to difficulty processing precise numerical details, particularly overlapping classes. It is important in preventing these errors that jurors comprehend the meaning of numerical amounts in context. The first step in communicating meaning to jurors is to identify the meaning of the numbers in evidence––the "bottom line" of the information. Expert knowledge is important in extracting this meaning. For example, an expert can tell jurors whether a particular number is low or high in a given context (e.g. whether a random match probability of 1 in 1000 is considered high in the field). This information can then be communicated to jurors alongside precise numerical values.

One way to effectively communicate the meaning of information in addition to precise numbers, while also avoiding the errors associated with focus on precise numbers, is to use visual aids and graphs.[151] These aids should be specifically designed to convey meaning and to avoid errors associated with processing––such as class-inclusion errors.[152] Visual aids such as column graphs, pie charts, 2x2 tables, and Venn diagrams can be used to clearly display probabilities and reduce processing interference.[153]

First, column graphs and pie charts can be used to make clear relative magnitudes.[154] Consider one of the biases described above based on how information is framed. Jurors are more likely to be convinced by DNA evidence when RMPs are presented as probabilities (e.g. a 0.1 chance of a random match) rather than frequencies (e.g. a 100 in 1,000 chance of a random match).[155] If this information is presented in column graphs or pie charts, it becomes clear that the two are equivalent to each other (see Figures 2 and 3 below). This technique has been successfully used to convey health related information to medical patients.[156]

---

151 *See* Wolfgang Gaissmaier, Odette Wegwarth, David Skopec, Anne-Sophie Müller, Sebastian Broschinski & Mary C. Politi, *Numbers Can Be Worth a Thousand Pictures: Individual Differences in Understanding Graphical and Numerical Representations of Health-Related Information,* 31 HEALTH PSYCHOL. 286 (2012) (noting that not all visual aids improve understanding).

152 *Id.*

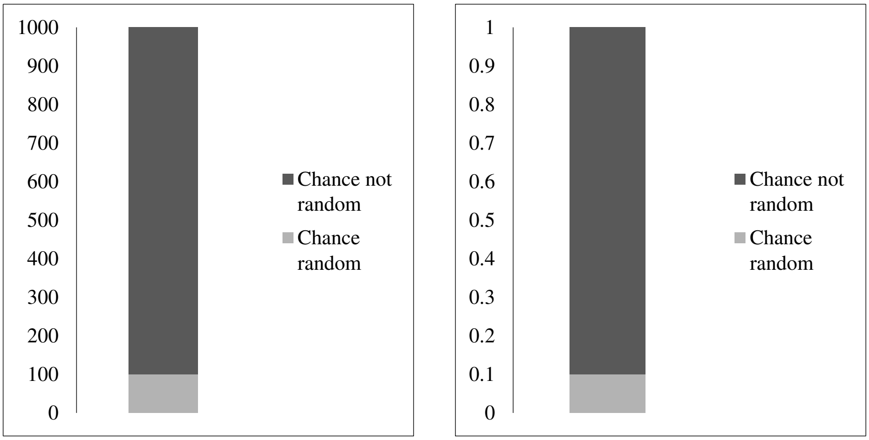
153 *Id.*

154 *See, e.g.* Liana Fraenkel, Ellen Peters, Peter Charpentier, Blair Olsen, Lanette Errante, Robert T. Schoen & Valerie Reyna, *A Decision Tool to Improve the Quality of Care in Rheumatoid Arthritis,* 64 ARTHRITIS CARE & RES. 977 (2012).

155 Koehler, *supra* note 31, at 502.

156 Fraenkel et al., *supra* note 154; Priscila G. Brust-Renck, Caisa E. Royer & Valerie F. Reyna, *Communicating Numerical Risk: Human Factors That Aid Understanding in Health Care,* 8 REVIEWS OF HUM. FACTORS & ERGONOMICS 235, 252–59 (2013).

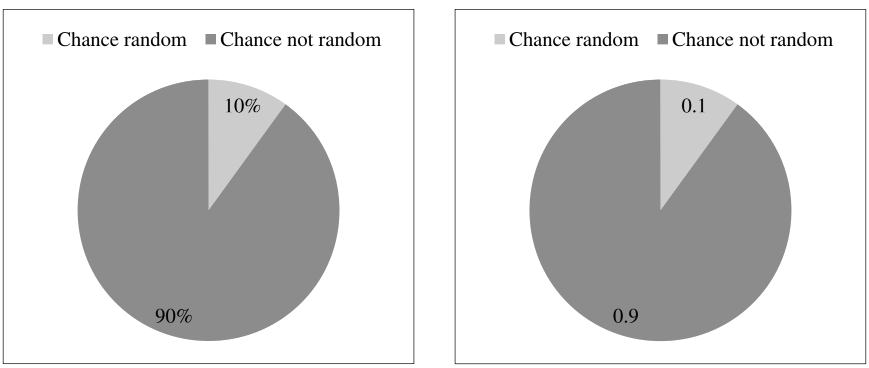
FIGURE 2. COLUMN GRAPHS SHOWING 100 IN 1,000 CHANCE AND 0.1 CHANCE OF A RANDOM DNA MATCH



Column graph displaying 100 in 1,000 chance

Column graph displaying 0.1 chance

FIGURE 3. PIE CHARTS SHOWING 100 IN 1,000 CHANCE AND 0.1 CHANCE OF A RANDOM DNA MATCH



Pie chart displaying 100 in 1,000 chance

Pie chart displaying .1 chance

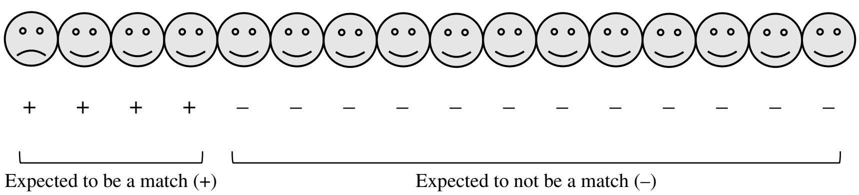
In addition to being good tools to communicate relative magnitudes, visual aids can be used to clearly separate classes from one another, in order to avoid class inclusion errors and to clearly communicate the meaning of statistics. Consider the following example (note that numbers used are not intended to be of realistic magnitudes but are included for illustrative purposes):

> *There is a group of sixteen suspects and it is thought that one of them committed a crime. A DNA trace is found at the scene of the crime and is a match to the defendant (who is one of the sixteen subjects). It is not known whether it is a match to the other suspects who are not available for testing. We know that there is a random*

> *match probability of 20%, meaning that there is a 20% chance of a match occurring by chance rather than because a defendant is the true source of a sample.*

In this example, many people would assume that this meant there was a 20% chance the defendant was innocent. This is an example of the prosecutor's fallacy discussed above, in which people equate the probability of the defendant being innocent given a match with the probability that there would be a match given the defendant is innocent (the random match probability). Using a visual aid like Figure 4 can help show why this is not correct, and make the correct probabilities evident. In Figure 4, the sad face is the true perpetrator, and the happy faces are innocent suspects.

FIGURE 4. VISUAL AID SHOWING NUMBER OF EXPECTED MATCHES IN SAMPLE OF 16 WHERE ONE OF THE SAMPLE IS GUILTY, AND THERE IS A RANDOM MATCH PROBABILITY OF 20%



Note. Sad face represents a guilty suspect; happy face represents an innocent suspect. Plus sign indicates a match; minus sign indicates not a match.

In this hypothetical, we expect the true source of the DNA trace (the one guilty person) to be a match, but we would also expect three of the remaining suspects to be a match since 20% of those who were not a true source would be expected to match (and 20% of fifteen is three). This means that given that the defendant is a match, there is a three-fourths chance that he is not the true source of the sample, and a one-fourth chance that he is the true source of the sample. Thus, even if we equate being the source of a DNA sample with being guilty, in this scenario there is only a one-fourth chance that the defendant is guilty. This technique can be used to help jurors in any cases where there is a risk of processing errors due to overlapping classes.

Although in real cases it is unlikely that there will be a pool of sixteen suspects, one of whom definitely committed the crime, it is important to estimate the probability that a defendant committed a crime based on diverse pieces of evidence, and to take this probability into account as one assesses the probative value of scientific evidence. In the case of DNA evidence, this would be an estimate of the probability that the defendant committed the crime, absent the DNA evidence. This could be high (because a lot of other evidence points to the defendant, or low (because the defendant was identified through a DNA database search). Considering the extent to which evidence other than DNA points towards guilt (the likelihood of guilt excluding the DNA evidence) is highly important in cases in which a defendant is on trial based on a DNA match from a database search, rather than due to other evidence.[157] This is because in these cases, setting aside the DNA evidence, the probability that the defendant left the sample is very small and thus the probability that a match is truly random is higher.[158]

For example, assume that the random match probability is 1 in 1,000. If a suspect is identified in a case due to other overwhelming evidence against her and then a DNA match is found, this is highly probative as to the likelihood of the suspect having committed the crime. If a database of 1,000 people is searched, then a match would be expected even if the person who was the true source of the sample was not in the database. Defendants have made arguments that DNA evidence obtained from database searches should be excluded on the basis that there is no generally accepted scientific opinion on how the probative value of such evidence should be quantified.[159] This issue has been considered by appellate courts, including the California Supreme Court, which decided that prosecutors may tell juries in all cases of the rarity of finding a defendant's DNA match, even when a database search has been used.[160] When evidence from database searches are employed, it is particularly important that processing errors are avoided, and jurors understand the statistics that they are evaluating.

One way to make the meaning of statistics here clear (and to avoid the prosecutor's fallacy and other processing errors) is to use visual aids, such as 2x2 tables, Venn diagrams, or stacked column charts.[161] Table 2 shows how a 2x2 table below can be used to display the different statistics involved.

---

157 *See* People v. Xiong, 215 Cal. App. 4th 1259 (2013).

158 *See* David J. Balding & Peter Donnelly, *Evaluating DNA Profile Evidence When the Suspect is Identified Through a Database Search,* 41 J. FORENSIC SCI. 603 (1996) (noting that although DNA evidence in such cases has been shown to be just as strong as DNA evidence from other cases, the suspect may be less likely to be guilty depending on the other evidence provided).

159 *See* David H. Kaye, *Rounding Up The Usual Suspects: A Legal and Logical Analysis of DNA Trawling Cases,* 87 N.C. L. REV. 425, 425 (2008).

160 People v. Nelson, 43 Cal. 4th 1242, 1267 (2008).

161 *See* Eric R. Stone, Winston R. Sieck, Benita E. Bull, J. Frank Yates, Stephanie C. Parks & Carolyn J. Rush, *Foreground: Background Salience: Explaining the Effects of Graphical Displays on Risk Avoidance*, 90 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 19 (2003).

TABLE 2. TWO-BY-TWO TABLE TEMPLATE DESIGNED TO PRESENT STATISTICS TO PREVENT THE PROSECUTOR?S FALLACY

	Probability of match (B)	Probability of no match (Not B)	Totals
Probability innocent (A)	Probability A and B	Probability A Not B	Probability of A =
Probability not innocent (Not A)	Probability B Not A	Probability Not A Not B	Probability Not A =
Totals	Probability of B =	Probability Not B =	Total = 100%.

Table 2 shows that the probability of a defendant being innocent given a match (Probability of A and B / Probability of B) is different from the probability of the defendant being a match given that she is innocent (Probability of A and B / Probability of A). Using such tables (particularly when populated with accurate numbers) can help to disentangle classes and make it easier for jurors to extract meaning from statistics. Tables 3 and 4 are examples of how such tables could be populated.

TABLE 3. TABLE DEPICTING A SITUATION IN WHICH THERE IS A PRE-EXISTING PROBABILITY OF 1 IN 1000 THAT A DEFENDANT COMMITTED A CRIME, A RANDOM MATCH PROBABILITY OF 1 IN 1000, AND A 0% CHANCE THAT THE PERSON WHO THE RECOVERED SAMPLE WAS FROM WOULD NOT BE A MATCH

	Chance of match	Chance of no match	Totals (based on existing evidence)
Person was true source of sample	**0.1 (100% of 0.1)**	0 (0% of 1)	0.1 (pre-existing chance defendant truly DNA source)
Person was not true source of sample (match was coincidental)	**0.1 (0.1% of 99.9)**	99.8 (99.9% of 99.9)	99.9 (pre-existing chance defendant was not the source of the DNA)
Totals	**0.2**	99.8	100

TABLE 4. TABLE DEPICTING A SITUATION IN WHICH THERE IS A PRE-EXISTING PROBABILITY OF 90 IN 100 THAT A DEFENDANT COMMITTED A CRIME, A RANDOM MATCH PROBABILITY OF 1 IN 1000, AND A 0% CHANCE THAT THE PERSON WHO THE RECOVERED SAMPLE WAS FROM WOULD NOT BE A MATCH

	Chance of match	Chance of no match	Totals (based on existing evidence)
Person was true source of sample	90 (100% of 90)	0 (0% of 90)	90 (pre-existing chance defendant truly DNA source)
Person was not true source of sample (match was coincidental)	0.01 (0.1% of 10)	9.99 (99.9% of 10)	10 (pre-existing chance defendant was not the source of the DNA)
Totals	90.01	9.99	100

Table 3 illustrates a case where a DNA match is found in a database of 1,000 people. Even if we know the real perpetrator is in the database, the pre-existing chance of a person from the database being the real perpetrator is 1 in 1,000. In contrast, Table 4 shows a case where a DNA match is found in a suspect, and independent evidence suggests there is around a 90% chance that she committed the crime. Of course, in real legal cases, jurors would be unlikely to have precise statistical estimates of guilt, but we use these figures and extreme examples to demonstrate potential contrasts. We also assume that if a person is the source of a DNA trace, he or she is the true perpetrator, although this would not necessarily be the case.

Note that in Table 3 (where a defendant was identified in a database search), even if there is a DNA match, the chance of the suspect being the true source of the DNA recovered from the crime scene and the chance of the match occurring by chance are equal (both 0.1/0.2). But where significant independent evidence suggested guilt, the chance of the defendant being the true source of the DNA recovered from the crime scene is much higher—around 99.99% (the chance of a match if the person was the true source which is 90 / the overall chance of a match which is 90.01). Putting this information in tables and explaining it to jurors can help jurors accurately extract meaning from numerical evidence, avoiding processing errors such as the prosecutor's fallacy.

D. *Providing Meaning Through Context*

Understanding the context of a number can also assist jurors in understanding the meaning of the number and avoiding biases associated with numerical decision making. In the context of DNA evidence, this could be done by presenting verbal equivalents alongside RMPs to show whether the chance of a random match is low or high in the context of DNA identifications. Experts can also provide context more generally when giving evidence, where this is not restricted by legal rules.

In the case of jurors translating meaning into numbers, research on meaningful anchors provides new insight into improving consistency.[162] Although jurors are prone to being influenced by meaningless anchor amounts, this research has shown that they are more influenced by meaningful anchors.[163] Meaningful anchors are numbers that are relevant to the case and to the decision regarding the award amount.[164] Providing these anchors for jurors is likely to increase consistency and decrease the influence of irrelevant anchors on jurors.[165] In the context of damage award amounts and contributory negligence judgments, one way to provide meaningful anchors may be to provide jurors with scaling instructions that help to consistently map a numerical amount onto a gist based judgment in an appropriate way.

The idea of using meaningful anchors to assist in judgment and decision making has been shown to be effective in other areas using "risk ladders."[166] These ladders compare one risk with another risk in order to increase risk understanding.[167] For example, in risk ladders that communicate environmental risks, exposure levels to chemicals and associated risk levels have been arrayed with low levels at the bottom and high levels at the top, helping people to anchor a risk to upper—and lower—bound reference points.

This work can be built on to develop "gist" scaling instructions that do not just scale numbers, but focus the jury on the few major qualifications that matter (e.g. in a pain and suffering case, the intensity of pain and suffering and the length of pain and suffering).[168] This should allow jurors to gain qualitative insights into the nature and magnitude of appropriate damages and avoid biasing influences, which should improve the ability to allocate appropriate responsibility to a plaintiff or to award an appropriate numerical damage award.

Damage award decisions and contributory negligence judgments require jurors to translate fuzzy qualitative concepts into numerical amounts. As noted above, this leaves much ambiguity, even when jurors

---

162 Reyna et al., *supra* note 73, at 285–87.

163 *Id.*

164 *Id.* at 280.

165 *Id.* at 288.

166 Carmen Keller, Michael Siegrist & Vivianne Visschers, *Effect of Risk Ladder Format on Risk Perception in High—and Low—Numerate Individuals,* 29 RISK ANALYSIS 1255, 1255 (2009).

167 *Id.*

168 DAVID BALL, DAVID BALL ON DAMAGES: A PLAINTIFF'S ATTORNEY'S GUIDE FOR PERSONAL INJURY AND WRONGFUL DEATH CASES (2d ed. 2005).

agree on qualitative aspects of a case. A gist scale building on previous work and focusing jurors on key meaning-based aspects of a case could help jurors map their perceived severity of injury and deservingness of damages onto a dollar scale, calibrating the severity of damage commensurately with an appropriate amount of money. In principle, scaling approaches can reduce irrelevant inconsistencies, such as larger damage awards for similarly-injured plaintiffs who sue richer (compared to poorer) defendants when actual damages are identical.

E. Active Jury Reforms

A final suggestion to help jurors understand and appropriately employ numerical information is to consider implementing active jury trial reforms. Some years ago, B. Michael Dann, then a trial judge in Arizona who regularly presided over jury trials, called for trial practice reforms that took into account jurors' active decision-making approach.[169] He observed that traditional jury trials were conducted with the apparent assumption that jurors were passive receptacles. In most jurisdictions even today, jurors cannot ask questions, nor are they able to talk with one another about the case until the final deliberations.[170] However, jury research shows that jurors take an active approach from the start of the trial, interpreting evidence and integrating it into a coherent narrative story.[171] What if they make errors in interpreting numerical information early on? Withholding the opportunity to clarify evidence and issues by forbidding juror questions or juror discussions during trial could be counterproductive, considering the active approach that jurors take to their decision making.[172]

There is now a significant amount of research on the best practices for lawyers and judges to implement active jury reforms, including question asking and trial discussions. The American Judicature Society's recommendations offer clear directions and outline protections for the

---

169 B. Michael Dann, *"Learning Lessons" and "Speaking Rights:" Creating Educated and Democratic Juries*, 68 IND. L.J. 1229, 1277–79 (1993).

170 GREGORY E. MIZE, PAULA HANNAFORD-AGOR & NICOLE L. WATERS, THE STATE-OF-THE-STATES SURVEY OF JURY IMPROVEMENT EFFORTS: A COMPENDIUM REPORT, 34–36 (2007) (describing frequency of trial innovations in state and federal courts as reported in lawyer and judge surveys); *see also* JURY TRIAL INNOVATIONS (G. Thomas Munsterman, Paula L. Hannford-Agor & G. Marc Whitehead eds., 2d ed. 2006) (identifying jury trial reforms and positive and negative claims about them).

171 DENNIS J. DEVINE, JURY DECISION MAKING: THE STATE OF THE SCIENCE 26–28 (2012) (explaining the story model of juror decision making).

172 *See* AMERICAN BAR ASSOCIATION, PRINCIPLES FOR JURIES AND JURY TRIALS, 17–19 (2005); Valerie P. Hans, *U.S. Jury Reform: The Active Jury and the Adversarial Ideal*, 21 ST. LOUIS U. PUB. L. REV. 85, 90 (2002).

parties.[173] When such reforms are implemented, jurors are generally enthusiastic about the value and benefits of being allowed to ask questions and being permitted to talk with other jurors during the trial under carefully specified circumstances.[174] Judges and lawyers tend to become more favorable toward these trial reforms as well, once they have experienced them.[175] We believe that trials with statistical evidence, cases that we know can prove challenging to jurors, are especially likely to benefit from the use of active trial reforms. Allowing juror questions during the trial and permitting jurors to speak with one another during the presentation of evidence could help jurors gain relevant background knowledge and appropriately critique statistical evidence. In addition, if jurors were encouraged to use 2x2 tables to write down statistical evidence, this could help avoid processing errors.

CONCLUSION

The *Apple v. Samsung* case may have been unusual in the amount of public scrutiny that it received, but it represents only the tip of the iceberg when considering jurors' evaluations of numbers. Jurors are frequently required to deal with numerical evidence, and to make judgments with a numerical value. As noted above, these tasks are difficult, and although jurors do a good job, they do often find these cases challenging. In this Article, we have explored some common sources of error and inconsistency in juror decision making, and have considered some of the primary reasons that these errors occur. Table 5 summarizes some techniques we have discussed that can be used to help jurors understand what specific numbers mean and thus improve juror dealings with numbers.

---

173 AMERICAN JUDICATURE SOCIETY, TOWARD MORE ACTIVE JURIES: TAKING NOTES AND ASKING QUESTIONS (1991).

174 Hans, *supra* note 172, at 92.

175 *Id.* at 92–93.

TABLE 5. SOURCES OF JUROR ERROR AND SUGGESTIONS TO REDUCE THESE SOURCES OF ERROR

Error	Suggestions to reduce error
Interpretation of evidence can be influenced by how information is framed	Use column graphs or pie charts to convey magnitudes, to show the equivalence of different values, and to reduce biases based on how values are framed.
Source probability error	Use 2x2 tables, Venn Diagrams, stacked column graphs, other visual aids, and active jury reforms.
Prosecutor?s fallacy	Use 2x2 tables, Venn Diagrams, stacked column graphs, other visual aids, and active jury reforms.
Difficulty combining two possible sources of error	Use 2x2 tables, Venn Diagrams, stacked column graphs, other visual aids, and active jury reforms.
Difficulty critiquing evidence given by experts	Provide statistical training for jurors, attorneys, and judges; use expert statisticians as a resource; use active jury reforms.
Inconsistency and unpredictability of jurors? numerical judgments	Use gist-based scales and scaling instructions.
Biasing influences on jurors? numerical judgments	Use gist-based scales and scaling instructions.

Although juror numeracy may help in some cases, it is not a cure-all, especially as many errors are likely to be caused by processing problems rather than a lack of knowledge. In addition, selecting only highly numerate jurors may actually introduce new systematic errors into juror decisions. Psychological research suggests that what is really important is understanding the meaning of numbers in context, something that can be done well by individuals who do not have a high level of technical numerical ability. This supports the role of the lay jury, and suggests that clear presentation of numbers and interventions to promote understanding of numbers and their meaning can help jurors dealing with numbers and ensure community values are reflected in judgments.