NUMERACY AND LEGAL DECISION MAKING

Arden Rowell* and Jessica Bregant†

ABSTRACT

This Article presents an empirical study of how numeracy—or math skill—relates to legal decision making. We describe three findings. First, the study shows a surprisingly high level of math skill among law students, especially given the common folk wisdom that lawyers are bad at math. Second, although prior research in non-legal contexts has shown that people with low numeracy are particularly susceptible to cognitive bias, we detect no significant relationship between law students' math skills and their susceptibility to bias or framing effects. Finally, and perhaps most strikingly, our findings show that the substance of legal analysis varies with math skill for at least some subset of cases. In particular, we find that law students with lower numeracy make decisions that are less consistent with negligence doctrine than students with higher numeracy.

---

*. Associate Professor and Richard W. and Marie L. Corman Scholar, University of Illinois College of Law. Thanks to Kenworthey Bilz, Dov Cohen, Mary Anne Franks, Valerie Hans, Paul Heald, David Hyman, Rob Kar, Bob Lawless, Jud Mathews, Lisa Milot, Anne Joseph O'Connell, Karen Bradshaw Shulz, and Jen Robbennolt for helpful comments and suggestions. Thanks also to Jeremy DeLaCerda for excellent research support, and to the Illinois Program on Law, Behavior, and Social Science for funding.

†. Ph.D. student, University of Chicago Booth School of Business; J.D., University of Illinois College of Law, 2009. Work for this project was completed while a Fellow in the Illinois Program on Law, Behavior and Social Science.

TABLE OF CONTENTS

INTRODUCTION

Law professors, judges, law students, and attorneys themselves routinely assume that lawyers are bad at math.[1] The assumption is so pervasive and casual that it has become a sort of in-group lawyer joke that attorneys tell to each other. Consider the words of First Lady Michelle Obama, addressing the National Science Foundation and explaining her choice of law as a career: "I know for me, I'm a lawyer because I was bad at [science and math]. All lawyers in the room, you know it's true. We can't add and subtract, so we argue."[2] Or consider a piece in the legal humor journal "The Green Bag," in which a practitioner facetiously speculated that the Green Bag's decision to charge $200 for a four-year subscription—in comparison to $40 for a year, $80 for two years, and $120 for three—might be "further evidence for the maxim that lawyers are bad at math."[3] The maxim is even used as an excuse when attorneys do make math errors: when IRS official Lois Lerner admitted to reporters that she was "not good at math[,]" after struggling to calculate one-quarter of 300, she tried to explain her difficulty by saying, "[b]ut I'm a lawyer, not an accountant, sorry!"[4]

This is not to say everyone jokes about lawyers' bad math skills; many commentators have also called gravely for math-based reform in the legal profession, as with Judge Posner's exhortation that lawyers and judges "overcome the prevalent (and disgraceful) math-block that afflicts the legal profession."[5] Yet judges and jokesters alike share the assumption that attorneys struggle with numbers.

---

1. For a discussion collating many disparaging remarks about lawyers' math skills, see Lisa Milot, *Illuminating Innumeracy*, 63 CASE W. RES. L. REV. 769, 776 (concluding that "when lawyers do math, they often do it badly"); *see also* MARY ANN GLENDON, A NATION UNDER LAWYERS: HOW THE CRISIS IN THE LEGAL PROFESSION IS TRANSFORMING AMERICAN SOCIETY 202–03 (1994).

2. *Remarks by the First Lady at the National Science Foundation Family-Friendly Policy Rollout*, WHITE HOUSE (Sept. 26, 2011, 4:06 PM), http://www.whitehouse.gov/the-press-office/2011/09/26/remarks-first-lady-national-science-foundation-family-friendly-policy-ro.

3. *See* Conor Moore, *To the Bag: Big Numbers*, 14 GREEN BAG 246 (2011). Moore also suggested that the journal might be "using its faithful readers as unwitting accomplices in a survey to determine if the maxim that lawyers are bad at math has any validity." *Id.* If true, the Green Bag has not reported on the results of its study.

4. Abby D. Phillip, *IRS Official: 'I'm Not Good at Math,'* ABC NEWS (May 10, 2013, 3:30 PM), http://abcnews.go.com/blogs/politics/2013/05/irs-official-im-not-good-at-math/.

5. *See* Richard A. Posner, *The Decline of Law as an Autonomous Discipline: 1962-1987*, 100 HARV. L. REV. 761, 778 (1987); *see also* GLENDON, *supra* note 1, at 202–03; Milot, *supra* note 1, at 809.

Why does it matter if lawyers *are* bad at math? One obvious and important concern is that poor math skills may lead judges, attorneys, and other decision makers to make math errors, which can affect the outcome of individual legal cases.[6] Even transparent math mistakes—such as adding damages wrong, or failing to correctly compound interest—should be taken seriously, as they can be catastrophic for the individuals involved.

These effects are reason enough to care about attorney numeracy. Yet recent literature in health and financial decision making suggests that low numeracy may have even more pernicious effects.[7] These literatures suggest that decision makers' numeracy relates not only to their ability to perform mathematical calculations; it is also related to the *substance* of their decisions. Or in other words, people with high numeracy make different decisions than people with low numeracy. And this is true even when the decisions are not obviously numerical:[8] performance on basic numeracy tests can be used to predict everything from whether a person is likely to seek preventive health screenings[9] or follow a prescribed medical regimen,[10] to how likely she is to be anxious about crime[11] or how likely she is to become unemployed.[12] Some studies even suggest that people with lower

6. *See* Milot, *supra* note 1, at 788–99 (chronicling the dangers associated with attorney math errors). Many of these errors may never be detected; nevertheless, between 2000 and 2007, there were 323 malpractice claims based primarily—if not solely—on "math error." *See* Dan Pinnington, *Avoiding Malpractice – Are you at Risk? The Most Common Legal Malpractice Claims by Type of Alleged Error*, July/Aug. 2010, at 29.

7. For overviews of these literatures, see NATHAN DIECKMANN, NUMERACY: A REVIEW OF THE LITERATURE, DECISION RESEARCH REPORT NO. 8-A, at 2 (2008); Peters et al., *infra* note 13, at 407.

8. For a useful and readable summary of this research in the medical context, see generally Valerie F. Reyna et al., *How Numeracy Influences Risk Comprehension and Medical Decision Making*, 135 PSYCHOL. BULL. 943 (2009).

9. *See* Lisa M. Schwartz et al., *The Role of Numeracy in Understanding the Benefit of Screening Mammography*, 127 ANN. INTERN. MED. 966, 966 (1997) ("[Higher] numeracy was strongly related to accurately gauging the benefit of mammography.").

10. *See* Kerri Cavanaugh et al., *Association of Numeracy and Diabetes Control*, 148 ANN. INTERN. MED. 737 (2008) (lower-numeracy persons were less successful in the self-management of diabetes); C. Estrada et al., *Literacy and numeracy skills and anticoagulation control*, 328 AM. J. MED. SCI. 88 (2004) (participants with lower numeracy engaged in poorer self-management of a medical regimen meant to prevent blood clots).

11. *See, e.g.*, Charles R. Berger, *Base-Rate Bingo: Ephemeral Effects of Population Data on Cognitive Responses, Apprehension, and Perceived Risk*, 29 COMM. RES. 99 (2002), *available at* http://rcirib.ir/articles/pdfs/cd1%5CIngenta_Sage_Articles_on_194_225_11_89/Ingenta824.pdf (measuring responses to burglary risk, and finding that participants lower in numeracy were more apprehensive about a reported increase in burglaries).

12. *See generally* SAMANTHA PARSONS & JOHN BYNNER, NAT'L RESEARCH & DEV. CENTRE FOR ADULT LITERACY & NUMERACY, DOES NUMERACY MATTER MORE? (2005) (finding a strong relationship between employment and numeracy, and comparing the impacts

numeracy are more subject to manipulation through cognitive biases and heuristics, because they are more likely to change their decisions based on how information is presented to them.[13] These effects add up, often to the cost of the innumerate: people with lower numeracy tend to live shorter lives,[14] and to be significantly sicker[15] and significantly poorer[16] than their more-numerate fellows.

If the substance of legal decisions—like the substance of medical and health decisions—varies with the math skill of the decision maker, the stakes are not small. At least on the margins, existing research suggests that we should be concerned that people may get different advice—and even different results—in identical legal cases, depending upon the numeracy of the attorney they employ or the numeracy of the judge (or jury) they face.

How piquant should this concern be? Up to this point, despite the widely held belief that attorneys are bad at math, and despite the increasing literatures on numeracy in finance and medical decision making, we can find no empirical study either of attorneys' basic numeracy level *or* of any potential interaction between numeracy and legal decision making.[17] Accordingly, this Article presents what we believe to be the first empirical study on the relationship between numeracy and legal decision making. In presenting this study and discussing the results, our goal is to begin to populate knowledge about the relationship between math skill and legal

of numeracy and literacy on employment; the study finds that numeracy is just as predictive as literacy for employment outcomes for men, and that numeracy is significantly more predictive than literacy for predicting women's employment outcomes).

13. *See* Ellen Peters et al., *Numeracy and Decision Making*, 17 PSYCHOL. SCI. 407, 412–13 (2006).

14. *See, e.g.*, David W. Baker et al., *Health Literacy and Mortality Among Elderly Persons*, 167 ARCH. INTERN. MED. 1503, 1503 (2007) (finding that lower health literacy—a measure based in part on numeracy—predicts shorter life expectancy).

15. *See, e.g.*, Valerie 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 (2007) (finding that people with lower numeracy make worse health decisions and have worse medical outcomes).

16. *See* James Banks & Zoë Oldfield, *Understanding Pensions: Cognitive Function, Numerical Ability and Retirement Saving*, 28 FISCAL STUD. 143, 143–44 (2007) (finding a strong relationship between numeracy and wealth, particularly post-retirement, even controlling for cognitive ability and education). *See generally* FINANCIAL LITERACY: IMPLICATIONS FOR RETIREMENT SECURITY AND THE FINANCIAL MARKETPLACE (Olivia S. Mitchell & Annamaria Lusardi eds., 2011) (discussing the relationship between wealth and financial literacy).

17. For the purposes of this discussion, we define legal decision making quite broadly, to mean decisions that are made by lawyers, in light of law, and/or that carry the force of law. In this sense, our search for relevant empirical work should, if anything, have been overinclusive.

decision making, and to explore the generalizability of nonlegal research to the specialized context of legal decision making.

Part I of this Article is meant to further acquaint the reader with existing research on numeracy and decision making, all of which is drawn from non-legal contexts. In that Part, we describe how numeracy is frequently measured, and how it appears to interact with non-legal decision making. Generally, this literature finds routine and predictable relationships between a decision maker's numeracy and the decisions she makes about her own health, wealth, and well-being. These relationships provide the motivation for the empirical study described in the latter parts of the article.

Part II presents an original empirical study. The study, which collects data from 150 law student participants, was designed to elicit three things: (1) participants' numeracy along the most commonly used numeracy measures, (2) a measure of participants' susceptibility to framing effects and cognitive bias, and (3) participants' substantive legal judgments in a simplified legal problem. In addition to illuminating underlying numeracy levels, this design allows us to compare law students' numeracy with their apparent susceptibility to bias, as well as to their substantive legal decision making. Among other findings, our study provides *no support* for the widely held belief that attorneys are bad at math. Second, although research in non-legal contexts has shown that low-numerate individuals have an increased susceptibility to framing effects and cognitive bias, we find *no* evidence that lower-numerate law students are more susceptible to cognitive bias than their more highly numerate colleagues, or at least not when they are acting within their zone of expertise. Finally, and strikingly, we do find strong evidence suggesting that substantive legal decision making can indeed vary with numeracy, at least for some types of legal questions.

Part III discusses some of the implications of these findings. Potentially important implications include the role numeracy plays in forming accurate legal predictions; the implications of numeracy for the future of legal scholarship and legal education; and the behavioral puzzle numeracy poses as to whether legal decision making is importantly different from decisions made in other life contexts. We also reflect on pathways for future research into numeracy and legal decision making.

I. NUMERACY: A REVIEW OF THE LITERATURE

Generally speaking, numeracy is the ability to understand and use numbers.[18] But what does it mean to "understand" and to "use" numbers? In a moment, we will describe some ways that researchers have answered this question. But before we get into those details, the reader might like to jot down the answers to these questions:

1) Imagine that we flip a fair coin 1,000 times. What is your best guess about how many times the coin would come up heads in 1,000 flips?
_____ *times out of 1,000*

2) In the BIG BUCKS LOTTERY, the chance of winning a $10 prize is 1%. What is your best guess about how many people would win a $10 prize if 1,000 people each buy a single ticket to BIG BUCKS?
_____ *person(s) out of 1,000*

3) In ACME PUBLISHING SWEEPSTAKES, the chance of winning a car is 1 in 1,000. What percent of tickets to ACME PUBLISHING SWEEPSTAKES win a car?
_____ *%*

The correct answers are footnoted.[19] The number of questions you answered correctly is your raw numeracy score on one widely used numeracy test, commonly called the "Schwartz test."[20] So if you got one

---

18. For an application of this definition, and for a very helpful review of the empirical literatures on numeracy, *see generally* DIECKMANN, *supra* note 7.

19. The correct answers are: 1) 500 2) 10 3) 0.1%.

20. For the first presentation of the test, see Schwartz et al., *supra* note 9 (using these questions to measure the numeracy of participants in a study on how people understand the benefits of mammography). For a discussion of uses of the Schwartz test, see DIECKMANN, *supra* note 7, at 11–12. As in many numeracy measures, participants in the Schwartz test have historically performed somewhat poorly: when it was first administered in 1997, only about half of participants were able to accurately state that a fair coin would come up heads 500 times in 1000 flips, and "one third of the sample thought that 1000 flips of a fair coin would result in <300 heads." Schwartz et al., *supra* note 9, at 969. Similarly, only half of participants were able to accurately identify the probable number of "Big Bucks" lottery winners as 10 persons out of 1,000, and only one in five was able to convert 1 in 1000 to 0.1%. *Id.* Overall, a full 30% of respondents failed to get a single question right, and only 16% of the participants answered all

answer correct, your score is a 1; if you got two correct, your score is a 2; and so on. Results on this test have been found to predict everything from your ability to gauge the expected benefits of a mammogram[21] to the amount of money you have saved for retirement.[22]

A number of numeracy researchers have relied on the Schwartz test, often supplementing it with additional questions.[23] The most influential supplement to the Schwartz test was created by Lipkus et al. in 2001.[24] It adds seven questions to the original three;[25] the subsequent combined set of ten questions has come to be known as the "Lipkus scale."[26] The scale was

---

three questions correctly (twenty-eight percent had a single correct answer, and twenty-six percent got two right). *Id.*

21. *See* Schwartz et al., *supra* note 9, at 966 ("[Higher] numeracy was strongly related to accurately gauging the benefit of mammography.").

22. *See generally* FINANCIAL LITERACY: IMPLICATIONS FOR RETIREMENT SECURITY AND THE FINANCIAL MARKETPLACE, *supra* note 16. People with higher numeracy have typically saved significantly more for retirement.

23. *See, e.g.*, Estrada et al., *supra* note 10 (asking participants to calculate the number of pills needed to equal a prescribed dose of a medication); Roxanne Parrott et al., *Risk Comprehension and Judgments of Statistical Evidentiary Appeals: When a Picture is Not Worth a Thousand Words*, 31 HUM. COMM. RES. 423 (2005) (supplementing the Schwartz test by asking participants to calculate the tip on a restaurant bill and to calculate 1/3 of 100, and finding that low-numeracy individuals were less persuaded by statistical evidence).

24. Isaac M. Lipkus, Greg Samsa & Barbara K. Rimer, *General Performance on a Numeracy Scale among Highly Educated Samples*, 21 MED. DECISION MAKING 37, 41–43 (2001).

25. *Id.* These questions, along with the Schwartz questions, make up the Lipkus test:

1) Which of the following numbers represents the biggest risk of getting a disease?

___ *1 in 100*, ___ *1 in 1000*, __ *1 in 10*

2) Which of the following numbers represents the biggest risk of getting a disease?

___ *1%*, __ *10%*, ___ *5%*

3) If Person A's chance of getting a disease is 1% in ten years, and person B's risk is double that of A's, what is B's risk? ___ *%*.

4) If Person A's chance of getting a disease is 1 in 100 in ten years, and person B's risk is double that of A's, what is B's risk? ___ *out of 100*.

5) If the chance of getting a disease is 10%, how many people would be expected to get the disease: A: *Out of 100?* ______. B: *Out of 1000?* ______.

6) If the chance of getting a disease is 20 out of 100, this would be the same as having a ____% chance of getting the disease.

7) The chance of getting a viral infection is .0005. Out of 10,000 people, about how many of them are expected to get infected? _____ *people*.

26. For a discussion of even more measurements of numeracy, see DIECKMANN, *supra* note 7, at 7–14. Although the Lipkus scale is now widely used, researchers continue to supplement the scale. *See, e.g.*, Ellen Peters et al., *Less is More in Presenting Quality*

originally used on a sample of highly educated participants to help disaggregate the effects of education and the effects of numeracy.[27]

The Schwartz and Lipkus scales remain the most widely used objective tests for numeracy used in the decision making literature.[28] But there are also influential measures for evaluating numeracy that are subjective rather than objective.[29] The most widely used of these is the Subjective Numeracy Scale ("SNS"), which was developed to address concerns that the objective scales were both stressful and time-consuming for participants.[30] The SNS relies on participants' descriptions of their own abilities and preferences in using numbers.[31] Thus, it asks participants to rate themselves in response to questions like "How good are you at working with fractions?" and "How

---

Information to Consumers, 64 MED. CARE RES. & REV. 169, 187–88 (2007) (adding four additional—and more challenging—questions to the Lipkus scale).

27. 87.9% of participants in the Lipkus study had greater than a high school education. *See* Lipkus et al., *supra* note 24, at 39 tbl.1. Although education is positively correlated with higher scores on numeracy tests, even among highly educated individuals, only 32% were able to accurately answer all of the additional questions on the expanded scale, and only 18% could correctly answer all of the questions on the 3-question Schwartz scale. *See id.*

28. More specific numeracy-related scales have also developed in some specialized decision making contexts. In the medical field, for example, "health literacy," or the set of basic skills that individuals must possess to function in today's healthcare system, is frequently thought to require basic numerical skills. *See* Ruth M. Parker et al., *The Test of Functional Health Literacy in Adults: A New Instrument for Measuring Patients' Literacy Skills*, 10 J. GEN. INTERNAL MED. 537 (1995). The Test of Functional Health Literacy in Adults, for example, includes a seventeen-item numerical ability section, along with a fifty-item section measuring reading comprehension. *See id.* Similarly, "financial literacy"—the set of basic skills that individuals must possess to make financial decisions in today's markets—encompasses a variety of numerical skills. Vanessa G. Perry & Marlene D. Morris, *Who Is in Control? The Role of Self-Perception, Knowledge, and Income in Explaining Consumer Financial Behavior*, 39 J. CONSUMER AFF. 299, 299 (2005) (including questions about personal savings rates, budgeting, and paying bills on time). Although these specialized measures are related to numeracy, both health literacy and financial literacy include skills and knowledge that are non-numerical, and that in many cases are specific to our current medical and financial worlds. For a bit more detail on these literatures, see DIECKMANN, *supra* note 7, at 9–10 (summarizing health and financial literacy and noting that "[b]oth health and financial literacy are relatively broad constructs that involve a range of skills including reading ability, domain specific knowledge, and numerical skills").

29. *See* Angela Fagerlin et al., *Measuring Numeracy Without a Math Test: Development of the Subjective Numeracy Scale*, 27 MED. DECISION MAKING 672, 672 (2007); *see also* Angela Fagerlin, Peter A. Ubel, Dylan Smith & Brian J. Zikmund-Fisher, *Making Numbers Matter: Present and Future Research in Risk Communication*, 31 AM. J. HEALTH BEHAV. S47, S48 (2007).

30. *See* Fagerlin et at., *Measuring Numeracy*, *supra* note 29, at 674.

31. *See id.* at 677 tbl.2.

good are you at calculating a 15% tip?"[32] This scale is significantly correlated with scores on objective scales, and is sometimes preferred because people report lower levels of anxiety when performing the test.[33] Relatedly, math confidence is sometimes used as a subjective numeracy-related measure.[34]

Scores on numeracy tests have been correlated with a number of demographic factors. Lower numeracy is associated with being female,[35] less educated,[36] African American,[37] Hispanic,[38] elderly,[39] and with having

---

32. *Id.* The Subjective Numeracy Scale asks four questions about what it calls "cognitive abilities"—questions 1 through 4 below—and four subsequent questions about "preference for display of numeric information." Here are the questions:

1. How good are you at working with fractions? *(1=not at all good, 6=extremely good)*

2. How good are you at working with percentages? *(1=not at all good, 6=extremely good)*

3. How good are you at calculating a 15% tip? *(1=not at all good, 6=extremely good)*

4. How good are you at figuring out how much a shirt will cost if it is 25% off? *(1=not at all good, 6=extremely good)*

5. When reading the newspaper, how helpful do you find tables and graphs that are parts of a story? *(1=not at all, 6 = extremely)*

6. When people tell you the chance of something happening, do you prefer that they use words (''it rarely happens'') or numbers ("there's a 1% chance")? *(1=always prefer words, 6=always prefer numbers)*

7. When you hear a weather forecast, do you prefer predictions using percentages (e.g., ''there will be a 20% chance of rain today'') or predictions using only words (e.g., ''there is a small chance of rain today'')? *(1=always prefer percentages, 6=always prefer words; reverse coded)*

8. How often do you find numerical information to be useful? *(1=never, 6=very often)*

33. *See id.* at 672 (finding a significant correlation between performance on the SNS and score on the Lipkus scale ($r = 0.68$)).

34. For a discussion on findings on math confidence, *see generally* Milot, *supra* note 1.

35. *See* Khaled Abdel-Kader et al., *Numeracy Skills in CKD: Correlates and Outcomes*, CLINICAL J. AM. SOC'Y NEPHROLOGY 1566, 1569 tbl.1 (2010); Mick P. Couper & Eleanor Singer, *The Role of Numeracy in Informed Consent for Surveys*, 4 J. EMPIRICAL RES. HUM. RES. ETHICS 17, 21 tbl.1 (2007) (numeracy measured using the Schwartz and Fagerlin scales); *but see* Ellen Peters & Erwin P. Levin, *Dissecting the risky-choice framing effect: Numeracy as an individual-difference factor in weighing risky and riskless options*, 3 JUDGMENT AND DECISION MAKING 435, 435 (2008) (finding no difference between male and female undergraduates' numeracy, as measured by the Lipkus scale).

36. *See* Peters & Levin, *supra* note 35. (finding that respondents who are college graduates are significantly more numerate than those with some college, who are significantly more numerate than those with a high school education or less).

37. *See* Abdel-Kader et al., *supra* note 35; Couper & Singer, *supra* note 35, at 21.

38. *See* Couper & Singer, *supra* note 35, at 21.

39. *See id.*

lower general mental ability.[40] These factors, however, by no means explain the range of numeracy described in the population: by one measure, demographic variations accounted only for one-fifth of the total variance in numeracy.[41] In other words, numeracy should not be conflated with any of these other characteristics—it is an individual status on its own, and one that can vary widely within otherwise similar groups.

II. NUMERACY AND LEGAL DECISION MAKING: AN ORIGINAL EMPIRICAL STUDY

Existing decision making literatures tell us that numeracy can play an important role in decision making about health, wealth, and personal well-being. But, how should we generalize these findings when the decisions under consideration are *legal* decisions—decisions made by people with legal training, which have the force of law and/or which address legal questions?

This Part presents an empirical studyevaluating the relationship between legal decision making and numeracy. The participants were law students, and the study measured both the students' numeracy and their legal analysis of several simplified legal problems. The study was intended to illuminate three puzzles: (1) whether attorneys—or people with legal training—are indeed "bad at math," (2) whether attorneys with lower numeracy tend to be more susceptible to manipulation of cognitive biases and framing effects, and (3) whether the math skill of a decision maker can affect the substance of legal decision making.

---

40. *See* Margaret Brooks & Shuang Yueh Pui, *Are Individual Differences in Numeracy Unique from General Mental Ability? A Closer Look at a Common Measure of Numeracy*, 8 INDIVIDUAL DIFFERENCES RES. 257, 262–63 (2010) (finding that numeracy—as measured by the Lipkus scale—is correlated with performance on a test for general mental ability, but concluding that the two measures tap into distinct processes).

41. *See* Couper & Singer, *supra* note 35, at 17 (accounting for age, education, race ("White," "Black," "Other"), Hispanic origin, gender, number of online surveys in the past month, and whether they presented the SNS or the Schwartz numeracy scale first). *See also* Brooks & Pui, *supra* note 40, at 263 (concluding that numeracy is distinct from general mental ability).

A. An Empirical Study of Numeracy and Legal Decision Making

1. Purpose

In running this study, we had three general research purposes. The first was to develop some empirical data regarding the numeracy of legal decision-makers that could be used to support or contend the general folk wisdom that attorneys are bad at math.[42] The second was to test the susceptibility of legal decision-makers to cognitive bias—and specifically, to determine the generalizability of existing findings that people with lower numeracy tend to be more subject to cognitive biases and framing effects.[43] Finally, this study sought to test the hypothesis that the numeracy of a decision maker might affect the substantive outcome of her legal judgments. This hypothesis was generated based upon the findings in existing numeracy literature that people with low and high numeracy tend to perceive and interpret the same facts differently, particularly when those facts relate to risks.[44]

We describe our methodology and results below.

2. Methodology

The study was performed using questionnaires that presented participants with a series of simulated legal problems—an experimental methodology.[45]

---

42. As we will explain, the study was run on law students. Law students are obviously not lawyers (yet). Given that legal education in general focuses very little on numerical skills, we see no particular reason that law students would not be good proxies for the attorneys they will become. That said, there appears to be a relationship between numeracy and aging, such that—at least once people edge towards retirement age—their numeracy tends to decrease. *See* Banks & Oldfield, *supra* note 16, at 151 (finding that, based on the English Longitudinal Study of Ageing and comparing people between ages 50 and 80+, "numeracy levels decline systematically with age"). Thus it could be the case that, because law students are younger (our average law student was 25 years old), their numeracy scores are on the whole higher than the (older) attorney population. There may be other important differences between law students and practicing attorneys, or between law students and judges, as well; to the extent these differences might affect numeracy, additional research on attorney and judge numeracy would be valuable.

43. *See, e.g.*, Peters & Levin, *supra* note 35, at 443–44 (finding that people with low numeracy are more loss averse and more sensitive to framing effects).

44. *See, e.g.*, Berger, *supra* note 11 (finding that the same information led to differential anxiety about crime, as mediated by numeracy); Schwartz et al., *supra* note 9 (finding that, while participants as a whole overestimated their risk of breast cancer, the effect was significantly exacerbated by low numeracy).

45. For a general overview of common types of empirical legal studies and their strengths and weaknesses, see Thomas S. Ulen, *Behavioral and Empirical Studies*, *in* THE OXFORD

a. Participants

Participants were 158 law students at the University of Illinois College of Law.[46] Overall, the study yielded 152 completed responses.[47] The study was only administered to students in required courses to limit potential selection bias.[48] Sixty-seven participants were students in the final week of their first year of law school who were enrolled in Constitutional Law, a required course at the University of Illinois. The remaining participants were second- and third-year students enrolled in Professional Responsibility, which is required under American Bar Association standards,[49] although students may choose when in their second or third years to take it.[50]

HANDBOOK OF BEHAVIORAL ECONOMICS AND LAW (Doron Teichman & Eyal Zamir eds., forthcoming 2014). The experimental methodology can be contrasted to a field study, which might have compared some measure of numeracy to real-world legal decisions, such as those made by practicing attorneys or judges. Field studies are often desirable because they incorporate real people working on real cases, but as any practicing attorney knows, no two cases are ever exactly the same. This means that the reality of field studies incorporates real variability across cases. Using an experimental methodology, instead of a field study, allows control of the information being presented to the decision maker, and allows us greater confidence that any difference in decision making comes from something other than slight and unmarked differences in the cases being considered. To see whether legal decision making varies with the numeracy of the decision maker, we need to hold as many other factors constant as possible. Thus an experimental study was the better fit for at least this initial study.

46. All research funded by the federal government and involving human subjects must be overseen by an Institutional Review Board. This study was approved by the Illinois Institutional Review Board. For an explanation of the relevant requirements, and a critique of current processes, see generally David Hyman, *Institutional Review Boards: Is This the Least Worst We Can Do?*, 101 NW. U. L. REV. 749 (2007).

47. Six surveys were omitted in their entirety because they were substantially incomplete and/or because the answers did not pass a basic check for sense.

48. The validity of a study can be threatened by selection effects "when an effect may be due to the difference between the kinds of people in one experimental group as opposed to another." *See* ROBERT M. LAWLESS, JENNIFER K. ROBBENNOLT & THOMAS S. ULEN, EMPIRICAL METHODS IN LAW 39 (2010) (summarizing additional potential threats to validity). If law students tend to select their courses differentially by numeracy, a law student population drawn from an elective course might not be representative of law student numeracy as a whole.

49. AM. BAR ASS'N, SECTION OF LEGAL EDUC. & ADMISSION TO THE BAR, 2012–2013 ABA STANDARDS AND RULES OF PROCEDURE FOR APPROVAL OF LAW SCHOOLS 19 (2012), *available at* http://www.americanbar.org/content/dam/aba/publications/misc/legal_education/Standards/2012_2013_aba_standards_and_rules.authcheckdam.pdf (Standard 302(a)(5)).

50. Two participants, one in a Constitutional Law class and one in a Professional Responsibility class, were enrolled in the Illinois LL.M. program, a Masters of Laws program for international lawyers. *See Apply LL.M*, UNIV. OF ILL. COLL. OF LAW, http://www.law.illinois.edu/prospective-students/apply-llm (last visited Nov. 18, 2013). These were included in the analysis, and did not affect the significance of any result.

Participants were 53.3% male, and their ages ranged from 21 to 40, with an average age of 25.0. Eight participants (5.3%) identified themselves as African-American, 12 (7.9%) as Asian, 15 (9.9%) as Hispanic, 111 (73.0%) as Caucasian, and 2 (1.3%) as "other."[51] As law students, all participants had at least a four-year college degree, and 10 participants (6.6%) indicated that they had also completed a Master's degree.

Participants completed the questionnaires during class time at the end of class. The professors in the course ended class a few minutes early, and students were then told by the authors—who were not teaching any of the relevant classes—that the students had an opportunity to participate in a study on decision making. The students were informed that the study was entirely anonymous, that their grades would not be affected if they chose not to participate, and that they were free to stop the study at any time. The students were not provided with anything in exchange for their agreement to participate.

There were no significant substantive differences between 1Ls and advanced students, or between different sections of the same class, so those distinctions are not discussed below.

b. Questionnaire

The questionnaire included three sections: measures of numeracy, questions that asked students to make a legal judgment or prediction, and demographic and non-legal questions. The demographic questions were always presented last, but the order of the other two sections was varied randomly. In addition, the order of questions within the numeracy section and the legal judgment section was varied randomly. The participants were randomly assigned to the various order conditions.[52] There was no significant effect of question order, so the responses were collapsed across order conditions for further analysis.

---

51. Participants were invited to select multiple racial categories if they identified as multi-racial. The University of Illinois College of Law reports its demographics as 394 (61.7%) Caucasian, 48 (7.5%) African American, 64 (10.0%) Asian, 45 (7.0%) Hispanic, and 267 (41.8%) female. *Enrollment Profile*, UNIV. OF ILL. COLL. OF LAW, http://www.law.illinois.edu/prospective-students/enrollment-profile (last visited Nov. 18, 2013) (data reflects full-time J.D. enrollment as of October 1, 2011).

52. This structure allows for between-subject comparison but not within-subject comparison. In other words, because each student only saw one version of the study, it was not possible to compare how the same student's answer might vary across different conditions.

c. Measuring Numeracy

Participants were tested for numeracy for two reasons: to provide some basis for establishing attorney numeracy, and so that a participant's numeracy could be compared to her answers to legal questions.

All of the participants completed the three-question Schwartz test.[53] The Schwartz test has been found to be highly correlated with performance on the other two frequently used numeracy measures: the longer objective Lipkus scale, as well as the Subjective Numeracy Scale (SNS).[54] To confirm that performance on these metrics was also comparable for law students, samples of students in the study were also given the full eleven-question Lipkus scale ($n = 38$), questions from the Subjective Numeracy Scale ($n = 39$), or both ($n = 23$). All participants also answered a subjective question measuring math confidence.[55]

d. Eliciting Substantive Legal Analysis

Participants were also presented with three simplified case vignettes and asked to rate the likelihood of various outcomes on a seven point, Likert-style scale.[56] These questions were designed to test different aspects of the general hypothesis that legal analysis performed by people with legal training varies according to the math skill of the analyzer. Borrowing from existing literature on numeracy and decision making, these questions sought to determine whether phenomena detectable in those studies also apply to legal decision making—i.e., decisions made about the law, on behalf of a third party, and/or by a decision maker with legal training.

Two of the questions, discussed in more detail below, were designed to measure the impacts of framing effects on legal decisions. The first question was designed to detect evidence of loss aversion and risk aversion.[57] It presented participants with either a win-oriented frame or a loss-oriented frame in the context of settlement negotiation. The second question was based on a question in the existing literature designed to measure

---

53. *See supra* note 19 and accompanying text for these questions.

54. *See supra* note 25 for these questions.

55. The question asked "How confident are you in your math skills?" Students were given seven options from which to choose: "not at all confident," "very unconfident," "somewhat unconfident," "neutral," "somewhat confident," "very confident," and "extremely confident."

56. Likert scale items, which are very common in psychometric research, ask participants to indicate the strength or intensity of their preferences or beliefs along a scale. *See, e.g.*, GUSTAV LEVINE & STANLEY PARKINSON, EXPERIMENTAL METHODS IN PSYCHOLOGY 368–72 (1994).

57. *See* discussion *infra* Part II.A.3.a–c.

susceptibility to probabilistic versus frequentistic framing effects.[58] It required participants to evaluate the legal risk posed by a potentially dangerous mental patient.

The third question shifted the focus from framing and cognitive bias to substantive legal prediction and judgment. It provided participants with numerical facts about a potential negligence case, and asked them to rate the likelihood of a negligence judgment.

3. Results & Discussion

a. Numeracy of Law Students

Out of the three questions of the Schwartz numeracy test, the law students answered an average of 2.41 correctly.[59] Their performances on each question are displayed in Table 1.

Table 1: Schwartz Test Performance

Question	Correct	Incorrect
Imagine that we flip a fair coin 1,000 times. What is your best guess about how many times the coin would come up heads in 1,000 flips? (500)	130 (85.5%)	22 (14.5%)
In the Big Bucks Lottery, the chance of winning a $10 prize is 1%. What is your best guess about how many people would win a $10 prize if 1000 people each buy a single ticket to Big Bucks? (10)	132 (86.8%)	20 (13.2%)
In Acme Publishing Sweepstakes, the chance of winning a car is 1 in 1,000. What percent of tickets to Acme Publishing Sweepstakes win a car? (0.1%)	105 (69.1%)	47 (30.9%)

A majority of the students (n = 87, 57.2%) answered all three questions correctly. Another 45 (29.6%) missed just one question out of three. Sixteen participants (10.5%) gave only one correct answer. Four students (2.6%) failed to answer any of the questions correctly.

58. *See* discussion *infra* Part II.A.3.d.

59. (*SD* = 0.78). In comparison, consider that the mean member of the public in the initial administration of the Schwartz test answered an average of only 1.28 questions correctly. *See supra* note 20.

As Table 2 shows, performance on the Schwartz test was strongly positively correlated with participants' reported math confidence, as well as to performance on the Lipkus test[60] and the subjective numeracy questions for the subsamples who completed these measures.

Table 2: Numeracy Measure Correlations

		Schwartz Test Score	Math Confidence	SNS Total Score	Lipkus Test Score
Schwartz Test Score	Correlation Sig. (2-tailed) *N*	1 152	0.22** 149	0.41** 39	0.45** 38
Math Confidence	Correlation Sig. (2-tailed) *N*		1 149	0.32* 39	0.14 37
SNS Total Score	Correlation Sig. (2-tailed) *N*			1 39	0.56** 23
Lipkus Test Score	Correlation Sig. (2-tailed) *N*				1 38

*** Correlation is significant at the 0.01 level.*
** Correlation is significant at the 0.05 level.*

This suggests that the short, three-question Schwartz test functions as a quick and easy measure of numeracy, at least as it would be measured by the longer Lipkus scale or the SNS. For this reason, the results reported here focus on numeracy as measured by the Schwartz test.[61]

Sixty-nine percent of participants reported that they are at least "somewhat confident" in their math skills. As Table 3 below shows, a majority of participants (56.4%) responded that they are "somewhat" or "very" confident. However, approximately one out of five respondents (20.1%) report being "somewhat unconfident" to "not at all confident" in their abilities.

60. For the subset (*n* = 38) who were given the full Lipkus test, mean score on the additional items was 5.16 (*SD* = 2.87) out of 8, and mean total score on the Lipkus test (which includes the Schwartz numeracy questions) was 7.45 (*SD* = 3.37) out of 11. For the subset (*n* = 39) who took the SNS, the mean reported subjective numeracy was 33.81 (*SD* = 7.60) out of 45 possible points.

61. *See infra* note 71. Researchers interested in examining the relationship between these measures are also welcome to contact us for our data.

Table 3: Math Confidence (Frequencies)

Response	***n***
Not at all Confident	6 (4.0%)
Very Unconfident	5 (3.4%)
Somewhat Unconfident	19 (12.8%)
Neutral	16 (10.7%)
Somewhat Confident	41 (27.5%)
Very Confident	43 (28.9%)
Extremely Confident	19 (12.8%)

Participants' math confidence responses were also coded as math confidence scores from zero to six, where a three represents the "neutral" response, and the mean math confidence for all participants was 3.92.[62] This corresponds to a mean response of "somewhat confident." Consistent with prior research on math attitudes,[63] men reported significantly higher levels of confidence than women.[64]

Past studies have found that lower objective numeracy is associated with being female,[65] African American,[66] and Hispanic.[67] With law students, this finding was not replicated as to Hispanic participants.[68] There were very

62. *SD* = 1.54.

63. *See, e.g.*, Janet Shibley Hyde et al., *Gender Comparisons of Mathematics Attitudes and Affect: A Meta-Analysis*, 14 PSYCHOL. WOMEN Q. 299 (1990).

64. Male participants ($M = 4.24$, $SD = 1.43$) reported significantly higher math confidence than female participants ($M = 3.59$, $SD = 1.57$), $t(140) = 2.59$, $p = 0.01$. The p-value tells us the chance that the result would occur randomly. A p-value less than 0.05 is typically considered significant in the social sciences; with a p-value of 0.05, we would expect the result to occur randomly only 1 time out of 20. A p-value of 0.01 is considered highly significant, as it is expected to occur randomly only 1 out of 100 times.

65. *See* Abdel-Kader et al., *supra* note 35, at 1568; Couper & Singer, *supra* note 35, at 17 tbl.1 (numeracy measured using the Schwartz and Fagerlin scales). *But see* Peters & Levine, *supra* note 35 (finding no difference between male and female undergraduates' numeracy, as measured by the Lipkus scale). *See supra* note 35.

66. *See* Abdel-Kader et al., *supra* note 35, at 1568; Couper & Singer, *supra* note 35, at 17 tbl.1 (numeracy measured using the Schwartz and Fagerlin scales); *see supra* note 35.

67. Couper & Singer, *supra* note 35, at 17 tbl.1 (numeracy measured using the Schwartz and Fagerlin scales); *see supra* note 35.

68. There was no significant difference between numeracy scores of students identifying as Hispanic ($M = 2.33$, $N = 15$, $SD = 0.82$) and those not so identifying ($M = 2.42$, $N = 137$, $SD = 0.78$), $t(150) = 0.42$, $p = 0.67$.

few students who identified as African American ($n = 8$), but those students did have moderately lower numeracy scores.[69] The mean numeracy score of men was also significantly higher than that of women.[70]

After calculating the numeracy scores of the participants, we divided them into two groups: the "high-numerate" group, who answered all three of the Schwartz test questions correctly, and the "lower-numerate" group, who answered at least one of the three questions incorrectly.[71] We did this for two reasons. First, because more than half of the law students answered all of the questions correctly, combining those who did not allowed for more balanced comparison and statistical analysis. Second, while more subtle analyses may be desirable in future studies, for this first study, comparing high- and lower-numeracy students gives us a clearer sense of the possible areas of distinction in legal reasoning and potential areas of concern. This simplification did not change the significance of any reported effect, and this categorization should, if anything, have made it more difficult to pick up numeracy-related effects. Thus, to the extent that it affects our results, we should be "under" reporting the effects of numeracy on legal decision making. That said, analysis of our results should take into account the fact that many "lower-numeracy" law students are in fact more numerate than much of the general population.

b. Likelihood of Settlement

The first legal question asked students about the decision to settle or go to trial. There were two versions of this question, and participants were

69. The mean numeracy score for African-American students was 1.88 ($SD = 0.84$), compared to 2.44 ($SD = 0.77$) for students who were not African-American, $t(150) = 2.02$, $p = 0.05$.

70. Men's ($M = 2.57$, $SD = 0.67$) and women's ($M = 2.28$, $SD = 0.83$) scores varied significantly, $t(143) = 2.31$, $p = 0.02$.

71. Researchers have used a variety of methods to identify or define low numeracy participants. *See, e.g.*, Woloshin et al., *Assessing Values for Health: Numeracy Matters*, 21 MED. DECIS. MAKING 382, 386 (2001) (identifying those who answer fewer than three out of three questions correctly as "low numeracy" participants). *But see* Peters et al., *supra* note 13, at 408 (using a median split to identify high and low numeracy). As we note, our sample was relatively evenly split between scores of three on the Schwartz test and scores less than three, so this split makes both intuitive and statistical sense.

It is also important to note that, of course, failing to answer one (or more) questions correctly on this measure does not necessarily indicate that person innumerate, or even low-numerate. For that reason, we have used the term "lower numerate" to describe these participants.

randomly assigned.[72] Although the versions are mathematically identical, one framed the outcome of the trial as a loss, and the other framed the outcome as a gain:

> *You are representing a client who is trying to decide whether to settle or go to trial. If your client prevails at trial, she will win $10,000. After researching, you determine that, for clients like yours, [1 out of 4 cases like this one are unsuccessful at trial / 3 out of 4 cases like this one are successful at trial]. Your client is offered a $5,000 settlement. Based on this information, how likely are you to counsel your client to settle?*

Participants were then prompted to choose their response along a seven-point likelihood scale.[73]

There were two hypotheses underlying this question, both based on prior research on numeracy and non-legal decision making.

The first hypothesis was that participants would exhibit risk aversion, but that this risk aversion would not vary by numeracy. Risk aversion arises where decision makers prefer certain choices to risky ones, even when the expected value of the risky option is greater than or equal to the value of the certainty.[74] In the context of this question, risk aversion would lead participants, attracted by the certainty of the settlement, to counsel their client to settle rather than take the risk of going to trial. In other contexts, researchers have found no interaction between numeracy and risk aversion: our hypothesis rested on the assumption that this finding would prove true in the legal context as well.[75]

---

72. Except where noted, participants were assigned to conditions randomly and independently, and conditions were fully-crossed. In other words, there was no relationship among the conditions to which a particular participant was assigned; each question was separately randomized. A total of 24 different versions of the written instrument were used.

73. As reported below, responses ranged from "extremely unlikely" (recorded as a zero) to "extremely likely" (recorded as a six). The midpoint of the scale was "neutral" (recorded as a three).

74. *See* Amos Tversky & Daniel Kahneman, *The Framing of Decisions and the Psychology of Choice*, 211 SCI. 453, 453–54 (1981).

75. *See* Pablo Brañas-Garza, Pablo Guillen & Rafael López del Paso, *Math Skills and Risk Attitudes*, 99 ECON. LETTERS 332, 336 (2008) (finding no differences in risk aversion across people who performed differently well on a GRE-like math test); Peters et al., *supra* note 13 (distinguishing between "risky choice framing effect" [i.e. risk aversion] and "positive" and "negative" frames [i.e. loss aversion] and finding that numeracy interacted with loss aversion but not risk aversion).

The second hypothesis related to loss aversion, which is the behavioral phenomenon whereby people tend to strongly prefer avoiding losses to acquiring gains.[76] Existing literature suggests that people with lower numeracy are more loss averse.[77] Thus, we hypothesized that there would be a greater difference in how likely students were to counsel settlement in the gain frame vs. the loss frame, as predicted by numeracy; participants with lower numeracy should be more affected by the shift from gain frame to loss frame.

Mean responses (on a scale from zero to six) for likelihood to counsel settlement by numeracy and condition are reported in Table 4.

Table 4: Mean (SD) Response by Numeracy

	Numeracy		
Condition	Low	High	**TOTAL**
3 out of 4 cases like this one are successful at trial	2.55[a] (1.29)	2.17[a] (1.24)	2.33[a] (1.27)
1 out of 4 cases like this one are unsuccessful at trial	3.70[b] (1.61)	4.02[b] (1.79)	3.88[b] (1.71)

Values with different superscripts differ at the $p < .05$ level.

These findings are notable in illuminating both risk aversion and loss aversion in legal decision making.

First let us consider risk aversion—or the preference for a certain choice over a risky one, even where the risky choice has a greater expected value. Past research has found that risk aversion is not meaningfully related to numeracy.[78] Our hypothesis was that this would hold true for legal decision making as well, and this hypothesis was supported by the data; there was no significant main effect of numeracy on likelihood of counseling settlement.[79] That said, our results did support the conclusion that law students are generally risk averse in the way that they would counsel clients

76. *See* Daniel Kahneman & Amos Tversky, *Choices, Values, and Frames*, 39 AM. PSYCHOLOGIST 341 (1984).

77. *See* Peters et al., *supra* note 13 (finding that numeracy interacted with loss aversion).

78. *See* Brañas-Garza et al., *supra* note 75, at 336 (finding no differences in risk aversion across people who performed differently well on a GRE-like math test); Peters et al., *supra* note 13 (distinguishing between "risky choice framing effect" [i.e. risk aversion] and "positive" and "negative" frames [i.e. loss aversion] and finding that numeracy interacted with loss aversion but not risk aversion).

79. $F(1, 146) = 0.01$, $p = 0.92$.

to settle: participants' average responses in the win-frame condition suggest that they were only "neutral" or "somewhat" unlikely to counsel their clients to settle; participants in the loss-frame condition were even less likely to counsel settlement; on average, they answered between "neutral" and "somewhat likely" to counsel settlement. The expected value of a $10,000 case with a 75% chance of success is $7,500. The settlement offer of $5,000 is thus significantly less than the expected value. Counseling the client to accept the offer is therefore risk averse, insofar as it counsels to accept a certain settlement in lieu of a risky (but probably more lucrative) trial.

Now let us consider these results as evidence of loss aversion. The results show strong support for the existence of loss aversion in law students' legal counsel.[80] Students were significantly more likely to counsel their clients to settle when the chances of success at trial were phrased in negative terms (the "loss" frame) than when the same chances were phrased in positive terms (the "gain" frame). More specifically, law students focused on the possibility of a loss ("1 out of 4 cases like this one are unsuccessful at trial") were significantly more likely to counsel their clients to settle than law students who were focused on the possibility of a gain ("3 out of 4 cases like this one are successful at trial"). This is true even though the underlying settlement offer ($5,000) and the underlying probabilities in the case are the same. This is a potentially worrisome finding, insofar as it suggests that substantive legal counsel may vary based upon loss aversion—and more specifically, upon whether an attorney ends up focused on the chance of success or the chance of failure.

Our results show that law students exhibited strong evidence of loss aversion. But did susceptibility to loss aversion in legal decision making vary with math skill, as it has been found to vary in the past with non-legal decision-making?[81] Apparently not: a two-way analysis of variance (ANOVA)[82] showed no significant interaction between condition (or whether the case was framed as a chance of a loss or a chance of a gain) and numeracy.[83] In other words, lower-numeracy participants did not show evidence of loss aversion at a significantly different rate than higher-

80. The main effect of condition—i.e., of whether the probabilities were presented in a gain frame or a loss frame—was highly significant. $F(1, 146) = 35.88$, $p < 0.001$, $\eta^2 = 0.20$.

81. *See* Peters et al., *supra* note 13 (finding that numeracy interacted with loss aversion).

82. An ANOVA (analysis of variance) evaluates whether the mean values of a given variable (in this case, perceived likelihood of negligence) are equal between two or more groups. A significant main effect indicates that, for at least one pair of means, there is a statistically significant difference.

83. $F(1, 146) = 1.97$, $p = 0.16$.

numeracy participants. Thus, this is an instance where existing literature on numeracy and decision making should not be generalized casually to legal decision making. Although legal decision makers are subject to loss aversion, they do not appear to be differently susceptible to it depending upon their math skills.

c. Likelihood of Counseling Release of a Patient

The next question was based on a question run by Peters et al., on medical professionals,[84] and asked the students how likely they would be to counsel release of a patient from a mental institution, given a numerical chance that he would be violent and a legal standard of unreasonable risk. This question also had two versions, varying only in how the probability of violence was expressed:

> *You represent a mental health institution where Mr. Jones is a patient. The institution is legally obligated to release Mr. Jones unless he poses an "unreasonable risk" to the public. According to the doctors at the clinic, of every 10 patients similar to Mr. Jones, [9/90%] commit no acts of violence to others during the first several months after discharge. Based on this information, how likely are you to counsel the institution to release Mr. Jones?*

Randomly-assigned participants in both versions were asked to rate their response along the same seven-point likelihood scale used in the previous question.

When a similar question was administered to medical professionals in the past, researchers found that lower-numerate individuals responded differently to the same information when it was presented in a probabilistic (e.g. "90%") frame than a frequentistic (e.g. "9 out of 10") frame.[85] Our version adds a legal standard and asks participants, who have some legal training, to apply the standard, a task meant to elicit legal as opposed to

84. Peters et al., *supra* note 13. Peters' question was very similar, though not identical: "Of every 100 patients similar to Mr. Jones, [10/10%] are estimated to commit an act of violence to others during the first several months after discharge." *Id.* Participants then rated the level of risk posed by Mr. Jones of harming someone on a scale ranging from 1 (low risk) to 6 (high risk). They found that, for highly numerate participants, the framing of the information was irrelevant, whereas individuals with low numeracy rated the risk much higher when it was presented in terms of frequency. *Id.*

85. *See id.*

medical decision making. Based on the Peters et al. study, we hypothesized that participants' decision to counsel to release Mr. Jones would be differently sensitive to the framing of the question, depending upon the numeracy of the participant. More specifically, and in line with Peters' findings, we predicted that lower-numerate people—and not higher-numerate people—would tend to think that Mr. Jones posed more of a risk when presented with the frequentistic frame, and less of a risk when presented with the probabilistic frame.

Because lower-numerate decision makers tend to experience a higher level of anxiety about risks,[86] we also hypothesized that lower-numerate law students might worry relatively more about the risk of releasing a potentially violent mental health patient. We therefore predicted that lower-numerate participants would be less likely to counsel release for Mr. Jones, irrespective of the version of the study they were analyzing.

Neither hypothesis was supported by our results. As to the first hypothesis, neither high- nor lower-numeracy participants showed any significant difference between conditions: a two-way ANOVA showed no significant effect of the interaction between framing condition and numeracy,[87] and no main effect of condition on likelihood of counseling release.[88] Thus, whether the risk was framed as a percentage or a ratio had no significant impact on whether participants of *any* numeracy counseled release. Moreover, there was no significant difference between the ways that higher- and lower-numeracy participants answered this question. This was striking in light of the findings of Peters, et al.,[89] and it is another instance where existing literature on numeracy and decision making should not be casually generalized to legal decision making.

The second hypothesis, that lower-numeracy participants would be more reluctant overall to release the potentially dangerous patient, was tested with a one-way ANOVA, and there was no significant difference between high- and lower-numeracy participants.[90] Thus, this hypothesis was also not supported, suggesting that something about the analysis makes participants'

---

86. *See* Berger, *supra* note 11, at 109–13 (measuring perceived burglary risk and attendant anxiety levels); Andrea D. Gurmankin et al., *The Effect of Numerical Statements of Risk on Trust and Comfort with Hypothetical Physician Risk Communication*, 24 MED. DECISION MAKING 265 (2004) (measuring perceived cancer risk).

87. $F(1, 145) = 0.59$, $p = 0.44$.

88. $F(1, 145) = 0.52$, $p = 0.47$.

89. *See* Peters et al., *supra* note 13.

90. Lower-numeracy participants' mean response ($M = 3.91$, $SD = 1.37$) did not differ significantly from that of high numeracy participants ($M = 3.69$, $SD = 1.57$), $F(1, 147) = 0.75$, $p = 0.39$.

decision making differ from analysis of risk in non-legal contexts.[91] In contrast with prior research, our results provide no evidence that lower-numeracy participants were more anxious about the risk posed by the patient. Because we did not attempt to measure anxiety directly, we cannot say whether the hypothetical simply failed to induce anxiety or whether the context of the question—or the position of the participant as a legal decision maker—somehow mediated the effects of the anxiety.

d. Substantive Legal Judgments: Probability of Negligence

This question, in contrast to the previous questions testing framing effects, was designed to evoke substantive legal judgment. It was based on the famous "Learned Hand formula," which holds that a tortfeasor is negligent if the burden of preventing a loss (B) is less than the probability of a loss occurring (P) multiplied by the magnitude of the loss (L), or if $B < P*L$.[92] This formula is arguably the most prominent approach used to determine negligence.[93]

Students were assigned[94] to one of four versions of the same question: *"Your client failed to take a $1,000 precaution that had a [1% / 5% / 10% / 20%] chance of preventing $5,000 in damages. The damage occurred. Based on this information, how likely is it that a court would find your client negligent?"* Reponses were given along a 7-point Likert-style scale ranging from "Extremely Unlikely" to "Extremely Likely," where the center

91. *Cf.* Berger, *supra* note 11, at 104–05 (finding lower-numeracy participants to be more concerned about burglary risk); Gurmankin et al., *supra* note 86, at 269 (finding lower-numeracy participants to be more concerned about cancer risk).

92. The formula is derived from language in United States v. Carroll Towing Co., 159 F.2d 169, 173 (2d Cir. 1947), which involved the question of whether a barge had been negligently moored. Learned Hand explained that "[s]ince there are occasions when every vessel will break from her moorings, and since, if she does, she becomes a menace to those about her; the owner's duty, as in other similar situations, to provide against resulting injuries is a function of three variables: (1) The probability that she will break away; (2) the gravity of the resulting injury, if she does; (3) the burden of adequate precautions. Possibly it serves to bring this notion into relief to state it in algebraic terms: if the probability be called P; the injury, L; and the burden, B; liability depends upon whether B is less than L multiplied by P: i.e., whether $B < PL$." *See also* The T.J. Hooper, 60 F.2d 737 (2d Cir. 1932). Judge Hand's choice to add an algebraic formula to his opinion may itself be an example of how numeracy and comfort with numbers can affect legal reasoning in meaningful ways.

93. *See* David Hunter & James Salzman, *Negligence in the Air: The Duty of Care in Climate Change Litigation*, 155 U. PA. L. REV. 1741, 1756–58 (2007).

94. For logistical reasons, the students assigned to the 20% condition were all drawn from the same 1L section. Because section assignments are themselves largely random, we have no reason to believe that this procedure added any sampling bias, and analysis of the demographics found none.

was "Neutral." Only one condition—the 20% condition—presented participants with a "balanced" scenario in which the burden of the precaution *equaled* the expected value of the precaution.[95] The remaining conditions, of course, suggested an even lower expected value. Thus, in no condition did the expected value of the precaution exceed its cost.[96] Given existing research suggesting that high numerate people are more likely to apply the rules of expected value,[97] we anticipated that individuals with high numeracy would tend to be more affected by the expected value calculation than those with lower numeracy. We therefore hypothesized that for the 1%, 5%, and 10% conditions, higher-numeracy participants would rate the probability of a finding of negligence as lower than participants with lower numeracy. We also expected that high-numeracy participants would draw greater distinctions between these conditions and the 20% condition, where the cost of the precaution was equal to the expected loss.

As we have discussed, all four conditions in this study involved an expected value that was less than or equal to the cost of the overlooked precaution, and the Hand Formula suggests that—at least on the basis of the limited information provided—the clients would likely not be considered negligent.[98] This also suggests an important distinction between the 20% condition, which is the "tipping point" in the Hand Formula analysis of this question, and the 1%, 5%, and 10% conditions, which we will call the "inefficient precautions" conditions. While legal doctrine is thus fairly clear that the client was not negligent in the three lower conditions, there is significantly less clarity in the 20% condition.[99]

Past non-legal numeracy research has found that individuals with lower numeracy tend to be subject to probability neglect, an insensitivity to

---

95. A 20% chance of successful prevention multiplied by the cost of the damages ($5,000) gives an expected value of $1,000, the exact cost of precaution in this scenario.

96. Due to the size limitations of our student sample, it would not have been feasible to include additional probability conditions in this question. Therefore, the question of how expected values that exceed the cost would have affected analysis was left for a forthcoming project.

97. *See* Daniel Benjamin, Sebastian Brown & Jesse Shapiro, Who is "Behavioral"? Cognitive Ability and Anomalous Preferences (May 5, 2006) (working paper), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=675264.

98. *See* discussion *supra* Part II.A.3.d.

99. Of course, this judgment assumes that the Hand Formula is the only relevant consideration. While this is undoubtedly a simplification of how negligence cases arise and proceed, we note that versions of the Hand Formula are often applied by courts as a definitive test of negligence. *See, e.g.*, McCarty v. Pheasant Run, Inc., 826 F.2d 1554, 1557 (7th Cir. 1987) ("Unreasonable conduct is merely the failure to take precautions that would generate greater benefits in avoiding accidents than the precautions would cost.").

changes in probabilities.[100] In the legal context, measurement of probability neglect is complicated by the fact that the legal impact of small probabilistic shifts can be discontinuous. Here, we expected that high-numeracy participants, predicting that the law would follow the Learned Hand formula—that is, participants resting their substantive decision making on the expected value of the precaution—would tend to show a sharp discontinuity between conditions where the precaution was inefficient (at 1%, 5%, and 10%), and the condition where the precaution was in equilibrium (at 20%). Lower-numeracy participants, who were expected to be less sensitive to changes in expected value, were expected to show *less* discontinuity between these condition sets. This distribution would look like probabilistic *sensitivity* if the observer were unaware of the discontinuous implications of the underlying legal doctrine.

The first hypothesis—that higher numerate participants would be more likely to apply the expected value calculation and therefore judge the likelihood of negligence as lower across the "inefficient precautions" conditions—was tested by comparing the responses of people with different levels of numeracy in each probability condition. Consistent with this hypothesis, lower-numeracy participants assessed a greater likelihood of negligence than higher-numeracy participants across the "inefficient precautions" conditions of 1%, 5%, and 10%, as shown in Figure 1.[101] At

---

100. *See, e.g.*, Benjamin et al., *supra* note 97, at 12–14; Couper & Singer, *supra* note 35, at 24 (explaining that lower numeracy is correlated with less sensitivity to extreme differences in disclosed risk, where numeracy was measured using the Schwartz and Fagerlin scales).

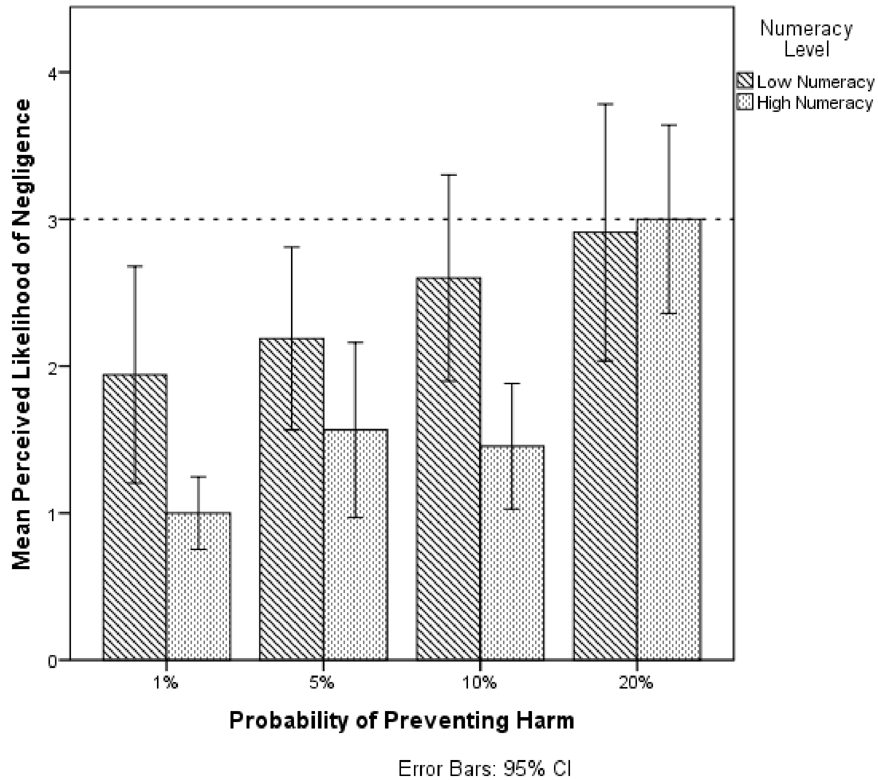
101. The differences between the higher- and lower-numeracy groups were statistically significant in the 1% condition, $t(42) = 3.01$, $p < 0.01$, and the 10% condition $t(40) = 2.97$, $p = 0.01$. In the 5% condition, the difference was in the predicted direction but not statistically significant. $t(37) = 1.48$, $p = 0.15$. Means and standard deviations are given in Table 5:

Table 5: Likelihood of Negligence by Condition and Numeracy

Condition	**Numeracy**	**Mean (SD)**	***N***
1%	Lower Numeracy	1.94 (1.44)	17
	High Numeracy	1.00 (0.62)	27
	TOTAL	*1.36 (1.10)*	*44*
5%	Lower Numeracy	2.19 (1.17)	16
	High Numeracy	1.57 (1.38)	23
	TOTAL	*1.82 (1.32)*	*39*

the 20% level, however, when the precaution had an expected value equal to the expected damage, there was no significant difference between the populations.[102]

Figure 1: Perceived Likelihood of Negligence by Condition and Numeracy



10%	Lower Numeracy	2.60	20
		(1.50)	
	High Numeracy	1.45	22
		(0.96)	
	TOTAL	*2.00*	*42*
		(1.36)	
20%	Lower Numeracy	2.91	11
		(1.30)	
	High Numeracy	3.00	14
		(1.11)	
	TOTAL	*2.96*	*25*
		(1.17)	

102. $t(23) = 0.19$, $p = 0.85$.

Consistent with our second hypothesis, lower-numeracy participants appeared to exhibit a lessened sensitivity to the expected value of the precaution. To find out whether this difference is significant, we collapsed across the "Inefficient Precaution" conditions and evaluated the role of numeracy and condition on perceived likelihood of negligence. The results are displayed in Figure 2 below.

Figure 2: Perceived Likelihood of Negligence by Efficiency Condition and Numeracy

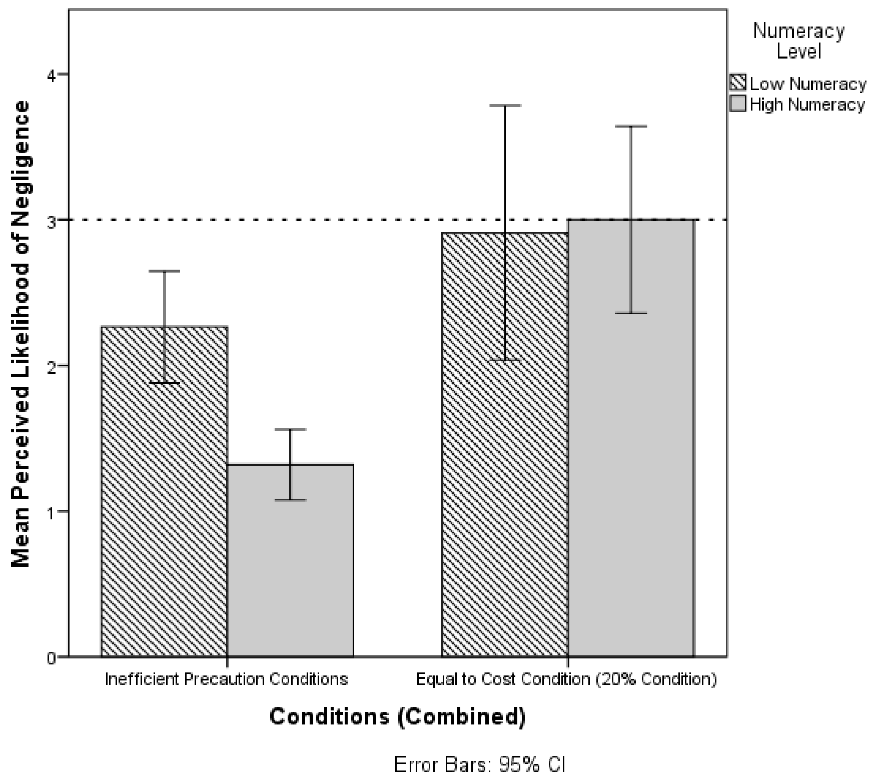


Figure 2 illustrates again that while there is little difference between participants in the equilibrium condition (20%),[103] higher-numeracy participants in the inefficient precaution conditions rated the likelihood of negligence as significantly less than did their lower-numeracy colleagues.[104] This analysis also confirms that the higher-numeracy participants drew a significant distinction between the equilibrium precaution and the

103. $F(1, 23) = 0.04$, $p = 0.85$.
104. $F(1, 123) = 19.05$, $p < .001$, $\eta^2 = 0.13$.

inefficient precautions, but that the lower-numeracy participants did not significantly distinguish between the two situations.[105] This is consistent with the hypothesis that for higher numeracy participants, the efficiency of the precaution played a greater role in their substantive determination of negligence.

These findings are consistent with our predictions about the doctrinal application. As shown above, high-numeracy participants judged the likelihood of negligence to be significantly higher in the 20% condition, and they did not significantly differentiate among the other three conditions. By contrast, lower-numeracy participants showed a gradually increasing perceived chance of liability, which increased with the chance that the precaution would be effective, but which showed no significant based on the efficiency of the precaution.

Although high- and lower-numeracy participants differed in their treatment of inefficient precautions, their answers in the 20% condition—where the cost of the precaution was equal to its expected value—were strikingly similar. Participants in both numeracy groups rated the likelihood of negligence in the 20% condition as essentially neutral, exactly as the Hand Formula would predict. For this condition, then, the substance of the participants' legal decision making did not vary by numeracy.

Why not? What makes the 20% condition different?

One possibility is that the 20% condition allowed participants to effectively ignore the numerical information once they had performed the expected value calculation. At that point, participants had to consider only the legal question: given that the cost and expected value were the same, how likely would a court be to find the client negligent? This is in contrast to the inefficient precaution conditions; in the 1%, 5%, and 10% conditions, participants would have had to judge the difference between the expected value and the cost—another number—and to judge the legal import of the numerical difference between the cost and the value.

The substance of participants' responses in the 20% condition is also consistent with doctrine; on our zero to six scale, 3.0 represents the "neutral" response, while answers of less than 3.0 mean that the participant

---

105. The interaction between numeracy and efficiency condition was marginally significant, meaning that the difference between lower-numeracy participants in the "Inefficient Precaution" conditions (M = 2.26, *SD* = 1.39) and lower numeracy participants in the "Equal to Cost" condition (M = 2.91, *SD* = 1.30) was less than the difference between higher numeracy participants in the "Inefficient Precaution" conditions (M = 1.32, *SD* = 1.03) and higher numeracy participants in the "Equal to Cost" condition (M = 3.00, *SD* = 1.11), $F(1, 146) = 3.84$, $p = 0.05$, $\eta^2 = 0.03$. We also did separate one-way ANOVAs for the higher-numeracy ($F(1, 84) = 30.34$, $p < .001$, $\eta^2 = 0.27$) and lower-numeracy ($F(1, 62) = 2.00$, $p = 0.16$) groups.

believes a negligence judgment is "somewhat" to "extremely" unlikely. As Figure 1 shows, while the average responses in the 20% condition were almost exactly three, the average responses for both numeracy groups in the 1%, 5%, and 10% conditions were less than neutral, squarely in the "unlikely" range. However, Figures 1 and 2 also show the averages were all significantly more than zero, or "extremely unlikely."

Given the doctrinal backdrop, why would participants—including highly numerate participants—say that there was *any* chance that a court would find a person negligent if she failed to engage in an inefficient precaution? Perhaps participants avoided committing to an extreme answer out of an abundance of caution, either because they are especially focused on the small chance that the client would be found negligent, because they recognized that real legal problems are rarely as clear-cut as the hypothetical in this survey, or because of a general extremeness aversion. Alternatively, participants may have believed that doctrine alone is an incomplete guide to legal prediction. For example, a highly sophisticated participant, aware that judgments of negligence may vary with the numeracy of the decision maker, might reasonably anticipate that a less-numerate judge or jury might find the client negligent even where the precaution was inefficient.

These possibilities, however, do not explain the differences between higher and lower numeracy participants. Prior research suggests that highly numerate people are more likely to apply expected value calculations.[106] If the expected value calculation in this question—operationalized through the Hand Formula—should lead participants to respond with lower likelihoods of negligence, then our results are consistent with lower numeracy participants applying expected value less strenuously.

4. General Discussion and Limitations

a. Numeracy of Law Students

This project started with the observation that attorneys are often described as "bad at math." The findings of the study presented here—the first we know attempting to directly measure the math skills of any group with legal training, including law students—suggests that this commonly held belief is untrue, or at least untrue for Illinois law students. A substantial majority (57.2%) of law students answered all three Schwartz

---

106. Benjamin et al., *supra* note 97, at 19.

questions correctly, a figure that contrasts starkly with the first reported administration of the test to the general population, where researchers found that just 16% of the sample was able to answer all three questions.[107] Even among Lipkus et al.'s "highly educated" participants, only 18% were able to answer all three Schwartz test questions correctly.[108] These findings suggest that attorneys are actually better at math than previous non-legal samples (at least insofar as attorneys are well-represented by law students). In comparison to most potential clients, then, or to jury pools, our findings do not support the claim that attorneys are innumerate.

Does this data suggest that critics are wrong to claim that attorneys are bad at math? Yes and no. At the very least, these findings suggest that it is misleading to categorically dismiss attorneys as "bad at math," especially as compared to the rest of the population. That said, our findings do support some potential criticisms of attorney numeracy. First, it may be that we miss something important about attorney numeracy if we focus only upon numeracy compared to the general population. Even if attorneys tend to be relatively good at math, perhaps they are still not good enough at math given the kind of responsibility that they often bear.[109] Attorneys are routinely put in significant circumstances of responsibility, and that responsibility may demand additional care when dealing with any question, numerical or otherwise.

Second, we have focused in our discussion on general tendencies. But individual attorneys' math skills may also be important. Through this lens, the fact that 3% of law students failed to answer a single numeracy question correctly may be quite disturbing. These students look to be innumerate: basically unable to manage even basic math problems.[110] If Illinois law students are representative of practicing attorneys, this would mean that approximately 37,000 lawyers licensed in the United States are functionally innumerate.[111] This is a worrisome back-of-the envelope calculation, and suggests that—even if most attorneys are likely to handle basic math problems perfectly well—we might still want systematic policy

107. Schwartz et al., *supra* note 9, at 969.

108. *See* Lipkus et al., *supra* note 24, at 37. Note that Lipkus' "highly educated" population included everyone who had at least "some college;" law students are obviously higher educated still, since they have all completed at least a four-year college degree.

109. *See* Milot, *supra* note 1 at 776.

110. For a valuable treatment of the implications of attorney innumeracy, see *id.*

111. According to the American Bar Association, there were 1,245,205 licensed lawyers in 2010. Three percent of this figure is 36,764. *See Lawyer Demographics*, AMERICAN BAR ASSOCIATION (2012), http://www.americanbar.org/content/dam/aba/migrated/marketresearch/PublicDocuments/lawyer_demographics_2012_revised.authcheckdam.pdf.

prescriptions to address the errors that creep into the system through those attorneys who are functionally innumerate.[112]

b. Law Students' Susceptibility to Cognitive Bias

An important finding in non-legal numeracy literatures was that decision makers with lower numeracy are more subject to cognitive bias and to framing effects, which can lead their decisions to be more easily affected by small changes in contextual cues.[113] If applicable to legal decision makers, this finding would have worrisome implications, as it would suggest that lower-numerate attorneys and judges might not only be more subject to math errors,[114] but that they also might be particularly subject to cognitive biases that could substantively change their analysis of legal questions.

Our findings do not support the hypothesis that lower-numerate persons with legal training are particularly susceptible to cognitive bias. We found no measurable effect of numeracy across multiple tests for risk aversion, loss aversion, or percentage/likelihood framing effects. This suggests that existing non-legal literature on the relationship between numeracy and bias should not be generalized to legal contexts, or at least not without significant further research.

Note that this finding—that people with law training who are asked to make legal decisions do not exhibit increased susceptibility to cognitive bias with lower numeracy—does not suggest that people with legal training are *immune* to cognitive bias. Literature on the operation of biases for experts remains limited, and results appear to vary over contexts and biases.[115] Insofar as law students qualify as "experts" for the purposes of this literature, our study makes two contributions in this area. First, it finds *no* evidence that law students are affected by probability/frequentistic framing effects when analyzing legal questions.[116] Second, and in contrast, we find strong evidence that law students *are* affected by loss aversion

112. *See* Milot, *supra* note 1, at 796 (discussing potential policy applications).

113. *See supra* discussion Part I.

114. As Milot worries, see Milot, *supra* note 1, at 780.

115. *See, e.g.*, Jeffrey R. Rachlinski, *Heuristics and Biases in the Courts: Ignorance or Adaptation?*, 79 OR. L. REV. 61, 63 (2000) (noting that experts are often just as likely to make cognitive and other mistakes as non-experts); *but see* Chris Guthrie & Jeffrey J. Rachlinski, *Insurers, Illusions of Judgment & Litigation*, 59 VAND. L. REV. 2017, 2025 (2006) (finding that expert insurers seemed to be less susceptible to heuristics than non-experts). Guthrie and Rachlinski argue that experience leads insurers to make more rational (i.e., unbiased) decisions; if they are correct, our law students are not "experts" at all, but highly-trained and highly-educated non-experts. *Id.* at 2047–48.

116. *See* discussion *supra* Part II.A.3.c.

when they are asked to make settlement decisions in the experimental context.[117] When students were directed to think about odds at trial in terms of success rate, for example, they were far less likely to counsel their fictional clients to accept an unfavorable settlement than if they were directed to think of the *same* odds in terms of failure rate. As stated above, however, neither of these findings varied by numeracy, or at least not for the relatively highly numerate population of law students.

c. Numeracy's Effect on the Substance of Legal Decisions

Another important finding in non-legal numeracy literatures was that decision makers' substantive decisions can vary with their numeracy.[118] Our findings are consistent with this aspect of the research. In other words, while susceptibility to cognitive biases does not seem to vary by numeracy among law students, substantive legal analysis *can* vary significantly with the numeracy of legal decision makers.[119] Does it matter if attorneys and other legal decision makers treat cases differently based upon their own math skills? We think it does, for several reasons.

First, attorneys often act as gatekeepers to legal action: a potential client who seeks advice as whether she should litigate a negligence action, and is told that her case is very weak, may never return to another attorney (perhaps one with better math skills) and may abandon the case. Furthermore, in many cases the actual legal outcome of a case is determined not by a judge, but by parties settling in consultation with their attorneys. A party may well accept—at the behest of her attorney—significantly less in settlement for a case deemed weak than one deemed strong: a difficulty if the apparent strength of the case is partially a function of the numeracy of the attorney. This means that attorneys have a substantive impact on the outcome of many—if not most—cases, even where there is no jury, judge, arbitrator, or mediator involved in the decision making.

That said, we do not mean to suggest that numeracy necessarily impacts every legal decision. In this initial study, we found only that determinations of negligence were related to the numeracy of the legal decision maker; further future analysis of additional types of legal decisions, and their relative relationship to numeracy, would be both welcome and helpful. In developing that future research, we think that our results should be taken to suggest that concern about differential legal perception, analysis, or advice

---

117. *See* discussion *supra* Part II.A.3.b.
118. *See, e.g.*, Gurmankin et. al, *supra* note 86.
119. *See* discussion of the Hand formula results, *supra* Part II.A.3.d.

may be particularly piquant where the legal issues in question directly involve risk or probabilities, as with determination of negligence.

d. Generalizability of the Study Population

We believe that this study presents valuable information about attorney numeracy. But what is the generalizability of our findings about law student numeracy to attorneys and other legal decision makers? We see two potential sources of concern in trying to generalize from Illinois law students' numeracy to the broader legal population.

The first potential concern is simply that law students tend to be younger than the attorneys, judges, arbitrators, and businesspeople that they will eventually become. Although we only have data for law students, we also worry that our findings suggest that other legal decision makers may be affected by numeracy. The law students in this study have proven to be surprisingly numerate, especially compared to previous research on the highly educated. Seasoned practitioners, who are—on the whole—older and more separated from their school days than law students, are likely to be somewhat less numerate.[120] Past research has found some relationship between numeracy and aging, such that—at least once people edge towards retirement age—their numeracy tends to decrease.[121] Because attorney numeracy has been studied so little, we can find no specific data showing whether attorneys' math skills also decrease with age, but if attorney numeracy follows the pattern of the general population, the law students we surveyed may tend to be better at math than attorneys nearing retirement.

Another potential concern arises from the fact that the students surveyed were all enrolled at the University of Illinois College of Law. Because numeracy has been so little studied, we have no specific data showing that students' choice of a law school relates to their numeracy. That said, we can see two potential factors that might tend to increase the numeracy of Illinois law students relative to students at other schools. The first is that the University of Illinois is a top-3 engineering and technology school. Law students with an engineering or technology background may be particularly

---

120. *See* Banks & Oldfield, *supra* note 16, at 143. Although our participants ranged in age from 21 to 40, the average was just 25 years old. Compare this, for example, to the average age of an active attorney in California—48—or the average age of a judge in California—60. *Member Demographics*, THE STATE BAR OF CALIFORNIA (Jan. 23, 2014, 11:00 PM), http://members.calbar.ca.gov/search/demographics.aspx.

121. *See* Banks & Oldfield, *supra* note 16 (finding that, based on the English Longitudinal Study of Ageing and comparing people between ages 50 and 80+, "numeracy levels decline systematically with age").

drawn to Illinois, given its general reputation in these areas. The College of Law has also had historical strengths in law and economics, corporate law, and empirical studies, any of which might similarly attract students and/or faculty with more quantitative background. Either of these factors might tend to increase the numeracy of Illinois law students' in comparison to the general population.

Overall, these factors might tend to drive up the numeracy of Illinois law students in comparison to the general legal population. At least until more research is done on numeracy and legal decision making, however, we should emphasize that the folk wisdom that attorneys are bad at math is not supported by empirical research. Rather, the best data we have at this point suggests that people with legal training actually tend to be quite good at math, particularly compared to the general public.

III. IMPLICATIONS AND DIRECTIONS FOR FURTHER RESEARCH

The study presented in this Article tested the proposition that math skill might matter to the substance of legal decision making. Our findings suggest that, at least for some subset of legal questions, substantive legal analysis can indeed vary with underlying math skill. Importantly, this demonstrates the necessity of thinking beyond arithmetic errors when considering the role of math in law. In the context of legal decisions, numeracy concerns also arise whenever decision makers most evaluate probabilities, risks, or calculations. These results hint at complex relationship between law and numeracy, and they raise a variety of continuing puzzles and questions for practitioners, researchers, and educators.

A. Accuracy in Legal Predictions

Does being bad at math make you bad at law? We find that law students predict different outcomes in identical cases, based on their own native math skills. But at least based on the evidence we have gathered, we would hesitate to draw any strong conclusions about whether this means that high- or lower-numeracy students are "better at law."

We can, of course, say that high-numeracy law students gave answers to this question that were more consistent with the Learned Hand doctrine. In this sense, they did "better" than their colleagues, and would likely score higher on a 1L exam. Insofar as the Learned Hand doctrine is an accurate predictor of how judges decide negligence cases, these students also did better than their colleagues at accurately predicting legal outcomes.

Before we conclude that higher-numeracy students (and attorneys) have an advantage at legal predictions, however, we should consider that the legal analyses of judges and juries may *also* vary with numeracy. If this is the case, the students who answered this question most accurately would be the ones who most accurately predicted the decision making of judges. If judges tend to have lower numeracy, then lower-numerate law students may well have been *more* accurate in their predictions than high-numeracy students.

Furthermore, our study was run on law students acting within their zone of relative expertise. But many legal decisions are made by laypersons, acting either as jury members or as clients. Because of this, the accuracy of lawyers' predictions about case outcomes relies not only on how attorneys and judges might view a case, but also on how clients and jurors view it.

Thus, the students who are most accurate at legal prediction will be determined by a function of the numeracy of the population of relevant judges and jury members—a question beyond the scope of this initial experimental study. Future studies on this question would be extremely valuable.

B. Math and Numeracy in Legal Education and Legal Scholarship

Our studies find that law students' substantive legal analyses vary with their underlying math skill, for at least some subset of legal questions. Here we sketch a few of the potential implications of this finding for legal education and legal scholarship.

Currently, prospective students in law school are not systematically screened for their numeracy, except incidentally, as through the "Analytical Reasoning" (often referred to as the "Games") portion of the Law School Admissions Test (LSAT).[122] Law schools do not typically require any level of mathematical education beyond that required to complete an undergraduate degree[123], and law schools do not typically require their

122. Law School Admission Council, *About the LSAT*, LSAC, http://www.lsac.org/jd/lsat/about-the-lsat (last visited Jan. 24, 2014) (analytical Reasoning questions test deductive reasoning abilities and the ability to understand a structure of relationships).

123. *See, e.g.*, *Admissions FAQ*, HARVARD LAW SCHOOL, www.law.harvard.edu/prospective/jd/apply/the-application-process/jdfaq.html (last visited Jan. 24, 2014) ("The Harvard Law School faculty prescribes no fixed requirements with respect to the content of pre-legal education."); *Eligibility*, YALE LAW SCHOOL, http://www.law.yale.edu/admissions/532.htm, (last visited Jan. 24, 2014) (applicants to Yale

students to take any math-related courses to complete their J.D., although a few law schools, such as NYU and Stanford, offer math-based courses such as "Accounting for Lawyers" or "Statistical Inference in Law."[124]

Our findings suggest that law schools should reconsider their general agnosticism towards their students' math skills. For good or ill, numeracy appears to affect students' substantive legal decisions. Law schools should respond to this knowledge at the least by affirmatively educating future attorneys to be aware of the impact their own math skills are likely to have on their analyses, and perhaps more aggressively by updating their curricula to include greater emphasis on the relationship between math skill and legal decision making.

Our findings also have potential implications for legal scholarship. The legal academy is currently undergoing what is often called an "empirical turn," and a significant subset of this work is increasingly quantitative.[125] The numeracy needed to process and evaluate much of this work is significantly above the minimal type of numeracy evaluation for which we were testing in our study, and presumably the authors of highly quantitative scholarship tend to have higher numeracy than those who do not do quantitative legal scholarship. Insofar as numeracy relates not only to the ability to use and understand numbers, but also to operations—such as risk perception and processing—that follow from that ability, we might expect the future to reveal an increasing chasm between high- and lower-numeracy academics, and the audiences they serve. Law faculties may experience increasing intra- or inter-faculty divergence not only on who is able to understand highly quantitative scholarship, but also on opinions about what constitutes good legal scholarship at all. And quantitative legal academics may experience an ever-increasing gap between their own numeracy—and the expectations and perception styles that go along with it—and the numeracy of laypeople, legislators, juries, and even attorneys and judges. At the least, this suggests that quantitative legal scholars should be aware that

---

Law School need only have received a bachelor's degree or its equivalent and have taken the LSAT).

124. *Accounting for Lawyers*, NYU LAW, https://its.law.nyu.edu/courses/description.cfm?id=11440 (last visited Jan. 24, 2014); *Statistical Inference in Law*, STANFORD LAW SCHOOL, http://www.law.stanford.edu/courses/empirical-evaluation-of-the-law (last visited Jan. 24, 2014).

125. *See* Mark Suchman & Elizabeth Mertz, *Toward a New Legal Empiricism: Empirical Legal Studies and New Legal Realism*, 6 ANN. REV. L. & SOC. SCI., 555, 556 (2010) (comparing the highly quantitative "empirical legal studies" movement with the more ecumenical "new legal realist" movement); *see, e.g.*, Gregory Shaffer & Tom Ginsburg, *The Empirical Turn in International Legal Scholarship*, 106 AM. J. OF INT'L LAW 1, 1 (2012).

their own numeracy-related decisions—e.g., to apply expected-value calculations when predicting legal outcomes—may or may not accurately predict the decision making of less numerically-facile decision makers.

C. Distinctiveness of Legal Decision Making and Legal Training

This Article has been concerned, in large part, with testing whether existing findings about numeracy and decision making from non-legal contexts apply to legal questions, and it has found that at least some of the findings from other contexts seem not to apply to legal decisions made by people trained in the law.

The question of whether a behavioral finding is generalizable to other contexts is a staple of social-science-based empirical analyses. The field of legal decision making, however, has yet to develop the habit of routinely querying whether "legal" decisions are importantly different from non-legal decisions in ways that matter to generalizability.

How is law distinctive (if it is)? Does it interact distinctively with existing ingredients? Or is it just one of many things that can be used interchangeably to thicken a soup? These are at least partly empirical questions, because they can be informed by research into whether legal decisions are made measurably differently than non-legal ones. But they are not purely empirical, because before we can attempt to observe potential differences in legal and non-legal decision making, we must have some concept of how to split up legal decision making from non-legal decision making—about what it is that we think could be importantly distinctive about law.

A thorough treatment of this issue is far beyond the scope of this article. At a first cut, however, we see at least three characteristic ways that legal decisions might plausibly and systematically vary from run-of-the-mill financial, consumer, or medical decision making.

The first characteristic is substantive. It could be that law itself is different than finance or health or consumption decisions, so that placing a decision in the "legal" bucket leads people to make different kinds of decisions than if they are faced with a "financial" question, or a question of health. The second characteristic has to do with third-party decision making. Legal decisions are often made by experts acting *on behalf of* someone else. This is different than most financial and consumer decisions, and it may trigger different kinds of decision making. And the third characteristic is that legal training to "think like a lawyer" might create distinctive decision-making pathways, such that legal experts may literally think differently even than other kinds of experts acting within their expertise.[126] Experts of any type may make different kinds of decisions than laypeople would with the same information and in the same circumstances.[127] But it could be that legal experts also think differently than other types of experts.

We have attempted to incorporate each of these qualities into the "legal" questions presented in our study. But further and more thorough analysis of this question would be valuable not only in future numeracy studies, but also in future analyses of any behavioral findings that are meant to transfer into legal contexts.

CONCLUSION

This Article presents the first empirical study regarding the relationship between numeracy and legal decision making. It concludes that, despite common wisdom to the contrary, people with legal training are surprisingly good at math. It also finds evidence that, while law students are susceptible to numerically-related cognitive biases in their legal decision making, their susceptibility does not vary significantly by numeracy. Finally, we find that the substance of legal decisions can vary with the numeracy level of the decision maker.

---

126. Experts (sometimes) appear to make decisions differently than laypeople do, at least when the experts are acting within the scope of their expertise. *See, e.g.*, Baruch Fischhoff et al., *Risk Perception and Communication*, 14 ANN. REV. OF PUB. HEALTH 183, 184 (2003)*; cf.* Chris Guthrie et al., *Inside the Judicial Mind*, 86 CORNELL L. REV. 778 (2001) (finding that federal judges were affected by five common heuristics and biases, although judges appeared less susceptible to two of these than laypeople).

127. *See* Fischhoff et al., *supra* note 126, at 184.