MATH SYMBOLS IN THE TAX CODE

WILL DANIELSON LANIER*

Our tax code is stuck in the Middle Ages. The Internal Revenue Code ("the Code"), codified at 26 U.S.C., uses the concepts of addition, subtraction, multiplication, and division, as one might expect of a tax code. But, disdaining the 1500s invention of the elementary math symbols '+,' '–,' '×,' and '÷,' the Code instead uses complicated English constructions such as "any amount of X which bears the same ratio to that amount as Y bears to Z."

I propose that we use these elementary math symbols in our tax laws. To see whether this would increase the laws' legibility, I conducted a preregistered, randomized, controlled trial involving 161 participants. One group received the actual Code, the other, a translation using math symbols. Both groups were asked to solve the same two Code-based tax problems. For the first problem, use of the translation with math symbols increased answer accuracy from 25% to 70%. For the second problem, answer accuracy increased from 11% to 50%.

This result, I argue, can be extrapolated to the broader population and to the Code as a whole, confirming the plausible intuition that math symbols would increase the understandability of the Code. I then argue that this would be a good thing, answering various objections along the way, with a particular appeal to the rule of law and the spirit of democracy. People ought to be able to understand the laws that govern them.

 LL.M., 2024, New York University School of Law; J.D., 2024, New York University School of Law; D.Phil., 2014, University of Oxford; B.Phil., 2009, University of Oxford; B.A., 2007, Cornell University. Thanks to Daniel Hemel, Samuel Estreicher, Caitlin Prentice, Francis Vergunst, and my family members for early comments, to Bob Beddor, Danny Foreman, Daniel Putnam, and Jack Samuel for helpful thoughts and encouragement, to my second-readers Leila Chang, Matthew Grossman, Colin Heath, Isabel Paolini, Ian Riley, Krishnan Sethumadhavan, and Yidi Wu, to my initial editor Adam Revello, to my primary editor Ian Riley, who has patiently put so much work into this piece, to all of the other editors at the *New York University Law Review*, and especially to Nora Danielson Lanier for edits and much more.

Introduction

In the year 1557, Robert Recorde, a Welsh doctor and book collector known for his "wide learning," published a book called *The Whetstone of Witte*.[1] It was a textbook on arithmetic written for the "great multitude, that desirously embrace all kindes of knowledge."[2] In it, Recorde first introduced to England the signs '+' and '–' to signify addition and subtraction.[3] Before this, English arithmetic books had used roundabout locutions, such as, "To whom it shal be addede 1, The nombre to be addede 1," instead of Recorde's more succinct "1+1."[4] Representing concepts in this way made mathematics easier, which furthered mathematical discoveries,[5] advanced science, and ultimately

---

1 Stephen Johnston, *Recorde, Robert*, Oxford Dictionary of National Biography (Jan. 3, 2008), https://www.oxforddnb.com/display/10.1093/ref:odnb/9780198614128.001.0001/odnb-9780198614128-e-23241 [https://perma.cc/T4D7-5NW3].

2 Robert Recorde, The Whetstone of Witte 2 (London, Jhon Kyngstone 1557), https://archive.org/details/TheWhetstoneOfWitte [https://perma.cc/SNR8-FLQS]. Recorde humbly acknowledged his limitations, but noted that "better it is that a simple coke doe prepare thy brekefast, then that thou shouldest goe a hungered to bedde." *Id.* at 3.

3 *See* Johnston, *supra* note 1.

4 Louis Charles Karpinski, The History of Arithmetic 103 (1925), https://catalog.hathitrust.org/Record/000166369 [https://perma.cc/28AG-HKQ4].

5 *See, e.g.*, Joseph Mazur, Enlightening Symbols: A Short History of Mathematical Notation and Its Hidden Powers, at xvi (2016) ("In mathematics, the symbolic form of a rhetorical statement is more than just convenient shorthand. First, it is not specific to any particular language . . . . Second, and perhaps most importantly, it helps the mind to transcend the ambiguities and misinterpretations . . . [and] to lift particular statements to their general form."); *id.* at 220 ("[Mathematical] ingenuities, simplification of complexities, making sensible connections–[are], in a large part, attributable to the illuminating efficiency of smart and tidy symbols."); A.N. Whitehead, An Introduction to Mathematics 59 (1911) ("By relieving the brain of all unnecessary work, a good notation sets it free to concentrate on more advanced problems, and in effect increases the mental power of [humanity]."). *See generally id.* at 58–70 (describing how notations enabled mathematical progress).

helped create the contemporary environment and lifestyles we see around us.[6]

Law, however, has lagged behind. When the first United States Congress passed the first federal tax statute, they declined to take advantage of Arabic numerals and other symbols, preferring instead medieval English (and even older Latin) words.[7] Ever since, the Internal Revenue Code ("the Code"), codified at 26 U.S.C., has employed lengthy, syntactically complex English phrases such as "any amount of X which bears the same ratio to that amount as Y bears to Z" instead of the elementary math symbols '+,' '–,' '×,' and '÷.'[8]

I propose that we use these elementary math symbols in our tax laws.[9] To see whether doing so would increase the laws' legibility, I conducted a preregistered, randomized, controlled trial involving 161 participants. The participants were randomly assigned to one of two groups: control or treatment. The control group received passages from the actual Internal Revenue Code; the treatment group received a translation of those passages using math symbols. Both groups were

---

[6] On the necessity of mathematics for science, see, for example, Eugene P. Wigner, *The Unreasonable Effectiveness of Mathematics in the Natural Sciences*, 13 Commc'ns on Pure & Applied Mathematics 6, 8–10 (1960) (explaining that "[t]he statement that the laws of nature are written in the language of mathematics was properly made three hundred years ago; it is now more true than ever before," and stepping through three concrete examples: planetary motion, quantum mechanics, and quantum electrodynamics). On the necessity of science for our built environment and lifestyles, see, for example, Vaclav Smil, How the World Really Works: The Science Behind How We Got Here and Where We're Going 67–88 (2022) (explaining the science behind "the four pillars of modern civilization": ammonia, steel, concrete, and plastics).

[7] *See* An Act for Laying a Duty on Goods, Wares, and Merchandises Imported into the United States, ch. 2, 1 Stat. 24, 26 (1789) (laying, for example, a tax on "all goods, wares and merchandises, other than teas, imported from China or India . . . twelve and a half per centum ad valorem").

[8] *See, e.g.*, 26 U.S.C. § 1(g)(3)(B) (mandating that a "child's share of any allocable parental tax of a parent shall be equal to an amount which bears the same ratio to the total allocable parental tax as the child's net unearned income bears to the aggregate net unearned income of all children of such parent to whom this . . . applies"). The drafters of the Code use the phrase "any amount [of X] which bears the same ratio to [that amount] as [Y] bears to [Z]" to mean "X×Y/Z." *See, e.g.*, I. Richard Gershon, A Student's Guide to the Internal Revenue Code 11 (4th ed. 1999) (explaining the phrase).

[9] Strangely, no one seems to have proposed this before. Many people have, however, argued that the Code should in general be simplified. *See, e.g.*, Clive Crook, *Want a Better IRS? Simplify the Tax Code*, Bloomberg (Aug. 19, 2022), https://www.bloomberg.com/opinion/articles/2022-08-19/want-a-better-irs-simplify-the-tax-code [https://perma.cc/V89U-QQ8C]. One article has argued that symbols from formal logic could help people understand the Code better—the article doesn't, however, advocate using the logic symbols in the Code itself, merely in one's efforts to understand the Code. Layman E. Allen & Gabriel Orechkoff, *Toward a More Systematic Drafting and Interpreting of the Internal Revenue Code: Expenses, Losses, and Bad Debts*, 25 U. Chi. L. Rev. 1 (1957).

then asked to solve the same two Code-based tax problems. For the first problem, use of the translation with math symbols increased answer accuracy from 25% to 70%. For the second problem, answer accuracy increased from 11% to 50%.

The texts and questions each group received are presented below,[10] and I urge you to try both the Code version and the mathematical translation to see whether you judge the mathematical symbols to be of help.

The experiment's result, I argue, can be extrapolated to the broader population[11] and to the Code as a whole,[12] confirming the plausible intuition that elementary math symbols would increase the understandability of the Code.

After explaining the results of the experiment, I discuss why this increase in understandability would be a good thing.[13] First, it would have good consequences.[14] Hundreds of millions of people, both in and out of the United States, collectively spend more than six billion hours each year trying to comply with the Code.[15] Disputes about their compliance result in over 70,000 lawsuits each year.[16] In many of these suits, nonspecialists, such as juries, pro se taxpayers, and United States District Court judges must apply the rules.[17] As a result of this whole process, the Code's inscrutability is both a costly time sink and a cause of demonstrably incorrect legal judgments.[18] And the Code's obtuseness siphons money away from low-income households into an industry that profits from, and lobbies for, the Code's complexity.[19]

Second, there are virtues of increasing understandability beyond what would show up in a cost-benefit analysis.[20] In particular, the rule of law and the spirit of democracy suggest that understandability is valuable in itself: All else equal, an easier to understand legal system is a better legal system.[21] And here all else really is equal; we can use math symbols to improve the understandability of our tax rules without changing their substantive content.[22] By doing so, we honor

10 *See infra* Section I.A.
11 *See infra* Section II.A.
12 *See infra* Section II.B.
13 *See infra* Section II.C.
14 *See infra* Section II.C.1.
15 *See infra* notes 92–93 and accompanying text.
16 *See infra* note 95 and accompanying text.
17 *See infra* note 96 and accompanying text.
18 *See infra* note 97 and accompanying text.
19 *See infra* notes 100–12 and accompanying text.
20 *See infra* Section II.C.2.
21 *See infra* notes 115–16 and accompanying text.
22 *See infra* notes 118–19 and accompanying text.

the principle that people ought to be able to understand the laws that govern them.

I
The Experiment

A. Methods

In August and September 2022, I preregistered this study,[23] and then I recruited 290 people online using Amazon Mechanical Turk, a common platform for social science research.[24] Each person was first given the following screener question to ensure that they were not a bot, and that they were trying to give the correct answer[25]:

> **Read the law below, and then answer the question that follows:**
>
> **The law:** Everyone shall receive a credit for half of his or her qualifying childcare expenses.
>
> **Question:** Mary has $200 of qualifying childcare expenses. **How much is her credit?**

---

[23] In the so-called "Replication Crisis" of the last decade, much previously lauded social science work has fallen into disrepute through bad experimental practices and a failure of the findings to replicate. *See, e.g.*, Andrew Gelman, *The Experiments Are Fascinating, but Nobody Can Repeat Them*, N.Y. Times (Nov. 19, 2018), https://www.nytimes.com/2018/11/19/science/science-research-fraud-reproducibility.html [https://perma.cc/R26M-A3SJ]; Edith Beerdsen, *Litigation Science After the Knowledge Crisis*, 106 Cornell L. Rev. 529, 530–32 (2021) (describing the history of the replication crisis). Because of Replication Crisis worries, before performing the experiment, I preregistered this study—filing on a public website the questions I would ask and the statistical methodology I would use to analyze the data, so that I could not change the methodology after-the-fact in order to manipulate undesirable data. *See Math Symbols in the Code*, OSF Registries (Aug. 28, 2022), https://osf.io/fa8d9 [https://perma.cc/9ZVN-B2AL]. For some background on the importance of preregistration, see generally Beerdsen, *supra* at 562–64 ("Preregistration has been called 'the most important outcome[] of the replication crisis' and described as 'the only way for authors to convincingly demonstrate that their key analyses were not *p*-hacked.'").

[24] *See Doing Academic Research with Amazon Mechanical Turk*, U.C. Berkeley Soc. Sci. Matrix (Oct. 14, 2021), https://matrix.berkeley.edu/research-article/doing-academic-research-with-amazon-mechanical-turk [https://perma.cc/64DP-RMN6] ("Amazon Mechanical Turk (MTurk) has become increasingly popular as an online tool for conducting social science research."). Mechanical Turk is a website by Amazon, where people can sign up either to perform jobs for others in exchange for money, or to post jobs that they will pay others to perform. *See generally Introduction to Amazon Mechanical Turk*, Amazon Web Servs., https://docs.aws.amazon.com/AWSMechTurk/latest/AWSMechanicalTurkGettingStartedGuide/SvcIntro.html [https://perma.cc/JS8M-LCFU].

[25] *See generally* Jon Agley, Yunyu Xiao, Rachael Nolan & Lilian Golzarri-Arroyo, *Quality Control Questions on Amazon's Mechanical Turk (MTurk): A Randomized Trial of Impact on the USAUDIT, PHQ-9, and GAD-7*, 54 Behav. Rsch. Methods 885, 886 (2022) (warning about reduced data quality from "inattentive workers, . . . intentionally dishonest workers, . . . 'bots,' . . . and virtual private networks"); *id.* at 885 ("These data suggest that the use, or lack thereof, of quality control questions in crowdsourced research may substantively affect findings . . . .").

The correct answer is, of course, half of $200, which is $100.

Out of the 290 participants, 161 participants gave the correct answer.[26] These 161 participants were randomly assigned to either the "Control" group or the "Treatment" group. Each member of each group got two tax questions to solve. Each tax question looked like the above screener question: It had a block of text set out as "the law," followed by a particular question prompt asking the participant to apply the law. The question prompts were the same for both groups. What varied between the groups was the text labeled "the law." For the Control group, the law text was a passage from the United States Internal Revenue Code, whereas, for the Treatment group, the law text was that same passage, but one I had partially translated using the math symbols '+,' '–,' '×,' '÷,' '=,' Arabic numerals, and variables. Each group entered their answer into a blank text box.

Here are the questions that the Control group received.[27] I encourage you, especially if you are not a tax expert, to get comfortable, grab a piece of scratch paper, and try the questions out on your own. See whether you find it easy or difficult to apply the Internal Revenue Code, and whether it becomes easier once you see the Code translated into mathematical symbols. Answers will be revealed at the end!

The Control group's two questions (I call the first the "Annuity Question," and the second the "Gain Question"):

> **Read the law below, and then answer the question that follows:**
>
> **The law:** Gross income does not include that part of any amount received as an annuity under an annuity, endowment, or life insurance contract which bears the same ratio to such amount as the investment in the contract (as of the annuity starting date) bears to the expected return under the contract (as of such date).

---

[26] 161 participants is a good sample size, by a standard means of deciding on the required size of a sample: statistical power. Power is, roughly speaking, an approximation of the probability that the study will find the hypothesized causal relationship, assuming that the causal relationship indeed exists. *See* DAVID C. HOWELL, STATISTICAL METHODS FOR PSYCHOLOGY 226 (7th ed. 2010) (describing "power, which is defined as the probability of correctly rejecting a false H_0 when a particular alternative hypothesis is true"). The a priori power of this study was 0.885, using the test for two proportions with different sample sizes, with effect size $h = 0.5$, $n_1 = 74$, $n_2 = 87$, $\alpha = 0.05$. Typically, studies aim for a power of no less than 0.8. *See id.* at 232, 241. To reach a power of 0.8, the study needed to have 63 people in each of the control and treatment groups (giving a total of 126 participants), using the test for two proportions with the same sample size, with effect size $h = 0.5$, $\alpha = 0.05$. *See generally* JACOB COHEN, STATISTICAL POWER ANALYSIS FOR THE BEHAVIORAL SCIENCES (2d ed. 1988) (describing the concept of statistical power and methods of calculating it).

[27] The passages labeled "The law" in the two questions for the Control group are direct quotes of 26 U.S.C. § 72(b)(1) and § 1001(a), respectively.

Question: David has received $100 as an annuity under a life insurance contract. The investment in the contract (as of the annuity starting date) is $5. The expected return under the contract (as of such date) is $10. **How much does gross income not include?**

Read the law below, and then answer the question that follows:

The law: The gain from the sale or other disposition of property shall be the excess of the amount realized therefrom over the adjusted basis provided in section 1011 for determining gain.

Question: Susan has a sale of property. The amount realized is $600. The adjusted basis provided in section 1011 for determining gain is $700. **How much is the gain from the sale of property?**

Well, how did it go? Once you are ready, I encourage you to keep reading, and try the questions again, this time with certain phrases of the Code replaced by elementary mathematical symbols.

Here are the Treatment group's versions of the two questions:

Read the law below, and then answer the question that follows:

The law: Gross income does not include the amount $a \times b \div c$

Where a is any amount received as an annuity under an annuity, endowment, or life insurance contract.

And b is the investment in the contract (as of the annuity starting date).

And c is the expected return under the contract (as of such date).

Question: David has received $100 as an annuity under a life insurance contract. The investment in the contract (as of the annuity starting date) is $5. The expected return under the contract (as of such date) is $10. **How much does gross income not include?**

Read the law below, and then answer the question that follows:

The law: The gain from the sale or other disposition of property = $a - b$, if $a - b$ is positive.

Otherwise, the gain = 0.

Here a is the amount realized.

And b is the adjusted basis provided in section 1011 for determining gain.

Question: Susan has a sale of property. The amount realized is $600. The adjusted basis provided in section 1011 for determining gain is $700. **How much is the gain from the sale of property?**

How did you find those questions, compared to the questions using the actual Code? For the Annuity Question, the correct answer is $50. For the Gain Question, the correct answer is $0.[28]

B. Results

For the Annuity Question, 87 participants were randomly assigned the actual Code language. Out of these, 22 (25%) got the correct answer.[29] 74 participants were randomly assigned the translation containing math symbols. Out of these, 52 (70%) got the correct answer.[30] The difference (25% versus 70%) is highly statistically significant.[31]

For the Gain Question, 87 participants were randomly assigned the actual Code language. Out of these, 10 (11%) got the correct answer.[32] 74 participants were randomly assigned the translation containing math symbols. Out of these, 37 (50%) got the correct answer.[33] The difference (11% versus 50%) is highly statistically significant.[34]

---

[28] Once each person completed the questionnaire, they received the New York City minimum wage ($15 per hour), *Minimum Wage*, NYC Business, https://nyc-business.nyc.gov/nycbusiness/description/wage-regulations-in-new-york-state [https://perma.cc/YP7B-Y9HR], for ten minutes of work, which is $15/6, or $2.50, regardless of how many questions they got right. They were aware of this compensation before starting the questionnaire, though the procedures of Amazon Mechanical Turk mean that experimenters can refuse to pay participants who do not put much effort into work. *See Approve and Reject Assignments*, Amazon Web Servs., https://docs.aws.amazon.com/AWSMechTurk/latest/RequesterUI/ApprovingandRejectingWork.html [https://perma.cc/97CQ-D9HM] ("When you approve an assignment, the Worker gets paid. When you reject an assignment, the Worker doesn't get paid."); *Approving and Rejecting Work*, Amazon Web Servs., https://docs.aws.amazon.com/AWSMechTurk/latest/AWSMechanicalTurkRequester/ApproveRejectWork.html [https://perma.cc/84M6-FSKT] ("We recommend you only reject work when workers are clearly putting in no effort to submit an accurate response to your task. It's inappropriate to penalize . . . worker[s] for submitting data incorrectly because you provided unclear instructions or they simply made a mistake in interpreting what you wanted them to do."). I did not reject any of the assignments. So, while every participant received the same amount of money regardless of how much time they spent on the assignment, they did not know that this would be the case. For all they knew, if they gave up on a problem too soon, they would not get paid.

[29] *See infra* Figure 1.

[30] *Id.*

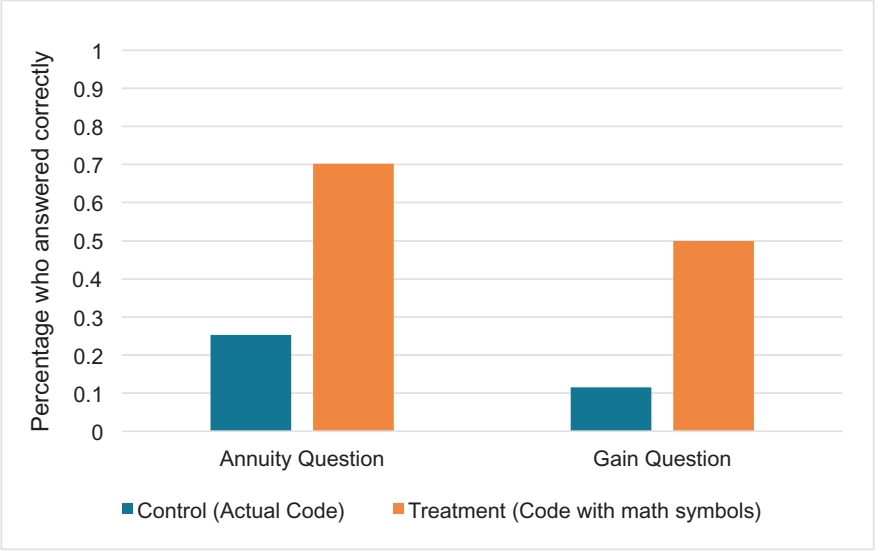
[31] $z = 5.7077, p = 0.00000001145$.

[32] *See infra* Figure 1.

[33] *Id.*

[34] $z = 5.3557, p = 0.00000008525$. As a comparison, I gave the test to ChatGPT shortly after it was released. This early version of ChatGPT answered the screener question correctly. When it was given the actual Code, it correctly answered the Annuity Question but not the Gain one. When it was given the translation in math symbols, the outcome was reversed: It correctly answered the Gain Question but not the Annuity one. *See* Interview with ChatGPT (Jan. 13, 2023) (on file with author).

FIGURE 1. USING MATH SYMBOLS IN THE CODE SIGNIFICANTLY INCREASED PROBLEM-SOLVING ACCURACY



For the Annuity Question, for Control, $n = 87$, of whom 22 (25.29%) got the correct answer. For Treatment, $n = 74$, of whom 52 (70.27%) got the correct answer. The difference between Control and Treatment is highly statistically significant ($z = 5.7077$, $p < .00001$). For the Gain Question, for Control, $n = 87$, of whom 10 (11.49%) got the correct answer. For Treatment, $n = 74$, of whom 37 (50.00%) got the correct answer. The difference between Control and Treatment is highly statistically significant ($z = 5.3557$, $p < .00001$).

II
THE UPSHOT

To draw conclusions from this experiment, I will break down the argument into several steps.

First, extrapolation[35]: I will argue that we can extrapolate from the random sample of these 161 participants to the U.S. population as a whole. If we tested the entire U.S. population on these two questions, we would likewise get a substantial increase in accuracy by using mathematical symbols in the Code. Furthermore, if people are better able to answer concrete questions using the law, then they understand the law better.

Second, the rest of the Code[36]: We can apply what we have learned from these two code sections to the Code as a whole. If these two laws

35 *See infra* Section II.A.
36 *See infra* Section II.B.

were small, unessential parts of the Code, or if they were some of the only provisions that would benefit from using math symbols in the Code, then using math symbols would not be of much help in understanding the tax law in general. On the contrary, however, I will show that the language of these two sections appears in many Code sections, including some of the most fundamental ones.

Third, the "So what?" question[37]: Even if using math symbols in the Code would make it more understandable, why should we care? I argue that there are important reasons why we should be trying to make the Code, and in general, the laws of the United States, more understandable to Americans (and to everyone). In this Section, I focus, inter alia, on the importance for the rule of law and democracy in having laws that are generally understandable.

A. Extrapolation

It would be quite disappointing if all we learn from this experiment are particular facts about 161 users of Amazon Mechanical Turk. In this Section, I argue that there are good reasons why the result concerning these users is at least roughly indicative of what the result would be if we asked all Americans.

Social scientists distinguish between internal and external validity. Say that a study claims that treatment X generally causes effect Y. The study has *internal validity* when, during the experiment, treatment X actually was the cause of the effect Y in the participants. If the study has internal validity, we have a further question: In what other circumstances does X cause Y? The extent of these circumstances is the question of *external validity*.[38]

When it comes to psychological research, such as this Article's study on the readability of texts, this is usually done on a particular subpopulation: college students.[39] Here, the question of external validity is particularly pressing, because psychologists are trying to learn about all humans, and there are clear differences between college students and the general human population.[40] Research done over the internet, as opposed to at a college campus, can improve upon this baseline,

---

[37] *See infra* Section II.C.

[38] *See* Francesco Guala, The Methodology of Experimental Economics 142 (2005) (giving a careful and extended definition of internal and external validity).

[39] Kenneth S. Bordens & Bruce B. Abbott, Research Design and Methods: A Process Approach 161 (9th ed. 2014).

[40] *Id.* ("If you use college students . . . the study will have less external validity. College students differ from the noncollege population in a number of ways (such as in age or socioeconomic status).").

because it can recruit from a more diverse participant pool, and let the participants be in a more familiar setting.[41]

Partly because of this, Mechanical Turk has been "rapidly adopted as a common research platform across the social sciences."[42] Many researchers believe that Mechanical Turk experiments "are at least as reliable as those obtained via traditional methods [and] are attractive for conducting internally and externally valid experiments."[43] Mechanical Turk experiments are "[w]idely used in social science research" because "samples of participants obtained via Amazon's Mechanical Turk (mTurk) tend to be representative across many sociodemographic variables."[44] For example, "results [have] suggested that American mTurk samples may be representative of the broader population in terms of global cognitive ability."[45] Experiments have found that Mechanical Turk results replicate many non-Turk results in a wide variety of fields: "decision-making, experimental economics, social psychology, cognitive psychology, clinical psychology, and political science."[46]

Nevertheless, some researchers disagree, and worry about the generalizability of Mechanical Turk experiments.[47] There are indeed

---

41 *See id.* at 162–63 (discussing studies comparing internet to non-internet results, and concluding that "the research in this area suggests that the Internet provides a powerful tool for researchers that may have fewer liabilities than critics allege").

42 Kyle A. Thomas & Scott Clifford, *Validity and Mechanical Turk: An Assessment of Exclusion Methods and Interactive Experiments*, 77 Computers in Hum. Behav. 184, 185 (2017). *See generally id.* (providing cites to many Mechanical Turk replications of previously well-established results).

43 Karoline Mortensen & Taylor L. Hughes, *Comparing Amazon's Mechanical Turk Platform to Conventional Data Collection Methods in the Health and Medical Research Literature*, 33 J. Gen. Internal Med. 533, 533 (2017). *See also* Thomas & Clifford, *supra* note 42, at 185 ("MTurk can provide researchers from a wide array of disciplines with data as reliable and valid as data collected in the lab . . . .").

44 Zachary C. Merz, John W. Lace & Alexander M. Eisenstein, *Examining Broad Intellectual Abilities Obtained Within an mTurk Internet Sample*, 41 Current Psych. 2241, 2241 (2022).

45 *Id.*

46 Thomas & Clifford, *supra* note 42, at 185 (cleaned up).

47 *See Doing Academic Research with Amazon Mechanical Turk*, *supra* note 24 ("'Demographically, this is not a representative sample of the US population, and you should never treat it that way,' Lenz said. 'If you're hoping to generalize your findings to the US population, don't.'"). The main worrisome studies involve questionnaires or surveys, for example, where the researcher asks if the participants use a particular social media service or know what an internet cookie is. *See* Jenny Tang, Eleanor Birrell & Ada Lerner, *Replication: How Well Do My Results Generalize Now? The External Validity of Online Privacy and Security Services*, Proc. of the Eighteenth Symp. on Usable Priv. and Sec. 367, 377 (2022), https://www.usenix.org/system/files/soups2022-tang.pdf [https://perma.cc/L8JS-NEA9] (discussing how, on these issues, results from Mechanical Turk can differ significantly from the general population). In other areas, Mechanical Turk experiments have been highly replicable. *See* Thomas & Clifford, *supra* note 42, at 185 (citing many replications using Mechanical Turk).

clear demographic differences between Mechanical Turk users and Americans as a whole. At the time of the experiment, about 90% of Amazon Mechanical Turk participants were in the United States.[48] Most of the rest were in India.[49] A recent study of the demographics of at least 1,000 U.S. and at least 1,000 Indian participants on Mechanical Turk found that, compared to the entire U.S. population, the U.S. Turk workers were, on average, people with "lower incomes[,] . . . younger[,] . . . more highly educated[,] . . . less likely to be married or report a religious affiliation[,] . . . [significantly less likely to be] Hispanic[] [or] African-American[, and] . . . disproportionately male."[50] The same study found that, compared to the Indian population as a whole, the Indian workers were "younger [and] . . . more educated[,] . . . [with a higher] monthly income," and that they were 67% male.[51]

Because of the demographic differences, it would be too quick to assume that, if the entire United States were studied, the percentages would be very close to the percentages found in the experiment. But one important thing to note is that the key factor is not the *percentage* of people who got each question right, but the *difference in percentages*

---

[48] *Countries*, MTurk Tracker, https://demographics.mturk-tracker.com/#/countries/all [https://perma.cc/BY28-GVKQ] (in date fields, enter "08/01/2022" and "09/30/2022").

[49] *Id.*

[50] Taylor C. Boas, Dino P. Christenson & David M. Glick, *Recruiting Large Online Samples in the United States and India: Facebook, Mechanical Turk, and Qualtrics*, 8 Pol. Sci. Rsch. & Methods 232, 236–37 (2020) [hereinafter Boas, Christenson & Glick, *Recruiting*]. The size of the differences between the U.S. averages and the mTurk averages for U.S. workers, respectively, were as follows: Age (49 vs. 34), Education (2.8 vs. 3.4, on a 5-point scale, where 2 is completed high school, 3 is some post-high-school but no bachelor's degree, and 4 is bachelor's degree), Income (5.8 vs. 5.2, on a 12-point scale, where 5 is total yearly family household income in the $40,000s, and 6 is in the $50,000s), Male (45% vs. 58%), Married (46% vs. 35%), Religious (79% vs. 47%), Hispanic (16% vs. 7%), Black (16% vs. 7%), White (75% vs. 85%). Taylor C. Boas, Dino P. Christenson & David M. Glick, *Appendix: Recruiting Large Online Samples in the United States and India: Facebook, Mechanical Turk, and Qualtrics* 22 tbl.3 (2020) [hereinafter Boas, Christenson & Glick, *Appendix*], https://www.cambridge.org/core/journals/political-science-research-and-methods/article/recruiting-large-online-samples-in-the-united-states-and-india-facebook-mechanical-turk-and-qualtrics/C80073966548D0E94161B84504ACE001#supplementary-materials [https://perma.cc/EM5Y-N5WV] (click on "Supplementary materials" and then "Appendix") (listing the information in the two columns labeled "Probability Sample" and "MTurk").

[51] Boas, Christenson & Glick, *Recruiting*, *supra* note 50, at 237. The size of the differences between the Indian averages and the mTurk averages for Indian workers, respectively, were as follows: Age range (2.95 vs. 2.1, on a 5-point scale which is not described), Education (2.8 vs. 6.4, on an 8-point scale, where 3 is Completed Class VIII but not Class X, and 6 is a bachelor's degree or equivalent), Income (4.7 vs. 7.0, on an 8-point scale, where 5 is total *monthly* household income of Rs. 4001–5000 (roughly, using the current purchasing power parity of $1:Rs24, this is $170–$210), and 7 is of Rs. 10,001–20,000 (roughly, $420–$830), Male (53% vs. 67%), Married (85% vs. 56%), Lower Caste (70% vs. 56%). Boas, Christenson & Glick, *Appendix* (listing the information in the two columns labeled "Probability Sample" and "MTurk").

between those who received the normal Code and those who received the version with math symbols. And for the demographic differences described above, it is not obvious that any of them shift the *difference* in percentages one way or the other.

For example, one would expect people with more education to be more likely to correctly answer the Treatment questions—the ones with mathematical symbols. But one would also expect them to be more likely to correctly answer the *Control* questions—the ones *without* mathematical symbols.

Math symbols can be confusing, but we tend to forget that English words are symbols too, especially since, as lawyers, we do so much reading.[52] The symbol '÷,' which I used in my translation of the Code, is a complicated one that my first-grade daughter has yet to learn—but then again, the same is true for the symbol 'ratio,' found in the untranslated Code above. In general, according to the Flesch-Kincaid Grade Level Test,[53] the text of the U.S. Code for the Annuity Question is Grade *Twenty-Three*, or, in other words, the Test suggests that one would need eleven more years of schooling after finishing high school to be able to comprehend the text.[54] The Code used in the Gain Question, untranslated, is rated by the Flesch-Kincaid Test to be Grade Seventeen.[55] In contrast, the mathematical symbols '+,' '–,' '×,' and '÷'

---

52 For a discussion of how tricky it is to keep in mind the distinction between the meaning of a symbol and the symbol itself, see Willard Van Orman Quine, Mathematical Logic 19–20 (rev. ed. 1981).

53 The Flesch-Kincaid Grade Level Test was developed by the U.S. Navy in order to estimate the number of grades of schooling one would need to have completed in order to understand a given text. *See Flesch Reading Ease and the Flesch Kincaid Grade Level*, Readable, https://readable.com/readability/flesch-reading-ease-flesch-kincaid-grade-level [https://perma.cc/29DN-L9US]. The test is admittedly crude, only depending on the number of words per sentence, and the number of syllables per word. *See id.* (stating the formula for the test). It is commonly used in legal research on readability. *See, e.g.*, Ian Gallacher, *"When Numbers Get Serious": A Study of Plain English Usage in Briefs Filed Before the New York Court of Appeals*, 46 Suffolk U. L. Rev. 451, 462–63 (2013) (using the Flesch-Kincaid Test to see changes in readability of briefs over time).

54 Photographs of Microsoft Word readability statistics (2023) (on file with author). I have used Microsoft Word to calculate the Flesch-Kincaid Grade Level Test scores for these texts. *See Get Your Document's Readability and Level Statistics*, Microsoft, https://support.microsoft.com/en-us/office/get-your-document-s-readability-and-level-statistics-85b4969e-e80a-4777-8dd3-f7fc3c8b3fd2 [https://perma.cc/9EJA-GC45] (describing how to use Word to calculate these scores). The scores are meant to reflect U.S. grades, and I note with amusement that, if we count law school and undergraduate degrees as a combined seven grades after high school, a graduating J.D. has only just finished Grade Nineteen, perhaps enabling them to be ready to read the Code section for the Annuity question after a further four years of training (say, a Tax LLM?).

55 Photographs of Microsoft Word readability statistics (2023) (on file with author).

are often introduced by Grade Three,[56] and the translated math-symbol versions of the Code from each question receive Flesch-Kincaid scores of Grade Eight and Grade Seven, respectively.[57]

There is a particular demographic difference, however, that could plausibly affect the difference in accuracy between those who received the normal Code and those who received the version with math symbols: English fluency.[58] As mentioned above,[59] about 10% of Mechanical Turk workers come from outside the United States. It is possible that Mechanical Turk workers outside the U.S. are at least slightly less fluent in English than Mechanical Turk workers inside the U.S., and if so, translating some of the English into math symbols could be more helpful to those living outside the U.S. Consequently, including them in the sample could make the difference in accuracy look larger than it would be if we just focused on the U.S. population.[60]

To check for this, I reran the analysis, excluding all those with an IP address outside the United States.[61] The differences observed did indeed shrink, but only slightly. For participants with a U.S. IP address, on the Annuity problem, use of the math-symbols translation increased answer accuracy from 27% to 69% (whereas when we include all IP addresses, the increase was from 25% to 70%).[62] On the Gain Problem,

---

56 *See Times Tables and More: In Third Grade, Multiplying and Dividing Take Center Stage*, Scholastic (2021), https://www.scholastic.com/parents/school-success/school-life/grade-by-grade/times-tables-and-more.html [https://perma.cc/W7TS-KVQH] ("Third grade is an important math year. Your child will learn how to multiply and divide . . . ."); Common Core State Standards Initiative, Common Core State Standards for Mathematics 21 (n.d.), https://corestandards.org/wp-content/uploads/2023/09/Math_Standards1.pdf [https://perma.cc/3TJF-YULD] ("In Grade 3 . . . Students develop an understanding of the meanings of multiplication and division . . . .").

57 Photographs of Microsoft Word readability statistics (2023) (on file with author).

58 Thanks to Adam Revello for pressing me on this.

59 *See supra* note 48 and accompanying text.

60 While I focus on the U.S. population, it is relevant to note that many people who are neither U.S. residents nor U.S. citizens still pay U.S. tax. *See, e.g.*, *Taxation of Nonresident Aliens – International Tax Gap Series*, IRS (Sep. 7, 2023), https://www.irs.gov/businesses/taxation-of-nonresident-aliens-international-tax-gap-series [https://perma.cc/84U7-42AF] ("Each year, thousands of nonresident aliens are gainfully employed in the United States. Thousands more own rental property or earn interest and/or dividends from U.S. investments.").

61 This would not exclude any who were using a VPN to obtain a U.S. IP address while being physically outside the country. Given that my survey allowed people with non-U.S. IP addresses to participate, however, such people would have no incentive, from my survey at least, to use such a VPN, though they could still have been using one for other purposes.

62 Out of 290 survey participants, 254 (88%) had a U.S. IP address, and 36 (12%) had a foreign IP address. The countries of the 36 foreign IP addresses were as follows: India (26), Brazil (3), Great Britain (3), Canada (1), France (1), Italy (1), and Spain (1). Excluding these foreign IPs, 254 participants remained, of whom 79 were assigned to Control and got the Screener question right. 21 of the 79 (so, 27%) also got the Annuity question right. 62 were assigned to Treatment and got the Screener question right. 43 of the 62 (so, 69%) also got the Annuity question right.

use of the math-symbols translation increased answer accuracy from 13% to 47% (whereas when we include all IP addresses, the increase was from 11% to 50%).[63] For both problems, the increases were again highly statistically significant.[64]

Why do the results shrink only slightly? Perhaps foreign Mechanical Turk users already have a high average fluency in English. Or perhaps it was because of the Screener question, which was in English, and only those who passed it were included in the experiment.[65]

Worries about external validity have a nice side benefit: They "suggest new problems that can be investigated experimentally."[66] Here, it would be nice to repeat this study among law school students, or among generalist judges or lawyers. These pools would be less diverse in some respects, but they would be interesting subpopulations to investigate. Also, another common strategy to address worries about external validity is to do gradual implementation in the "field," with lots of testing.[67] For example, with drugs, researchers conduct animal experiments, then small human experiments, then "efficacy trials" with humans in more representative settings.[68] Similarly, with math symbols in the tax code, we could try using them in a single new law, and then gather data about the effects.

In sum, I think we can conclude that, while the actual percentage of people who got each question right would probably decrease if we got answers from everyone in the United States, it is not obvious that the *difference* between the percentages would significantly decrease or increase. True, we should not be very confident about the exact amount of the difference. But given that the differences found in the experiment are *very large*—roughly, a three-fold increase in accuracy in the Annuity

---

63 Excluding the foreign IPs, on the Gain question, in the Control group, 10 of the 79 (so, 13%) got it right. In the Treatment group, 29 of the 62 (so, 47%) got it right.

64 For the Annuity question, $z = 5.0634$, $p = 0.0000004$. For the Gain question, $z = 4.4951$, $p = .0000007$.

65 Again, the Screener question was: "Everyone shall receive a credit for half of his or her qualifying childcare expenses. Mary has $200 of qualifying childcare expenses. How much is her credit?" *See supra* note 25 and accompanying text. As an analogy, here's a German translation of the Screener question: "Jeder erhält einen Kredit für die Hälfte seiner oder ihrer berechtigten Kinderbetreuungskosten. Maria hat $200 an berechtigten Kinderbetreuungskosten. Wie viel ist ihr Kredit?" This would certainly screen me out of a German survey.

66 GUALA, *supra* note 38, at 157.

67 *Id.* at 198.

68 *Id.*

question and a four-fold increase in the Gain Question—we should expect that, in general, people would find the math-symbol versions significantly easier to understand than the original Code versions. This has certainly been my experience anecdotally when presenting people face-to-face with the versions. I also urge the reader to consider your own experience of looking at the different versions of the questions. Intuitively, the experiment's general conclusion—that the math symbols translation is more understandable than the untranslated text—would be robust to changes in readers' characteristics. Looking back at the two questions, one would be hard-pressed to think of a group of people who would do *better* on the untranslated than the translated version.

B. The Rest of the Code

The experiment as such concerns only two short passages from Title 26 of the United States Code, which in total contains more than one million words.[69] If these two short passages were inessential, and unrepresentative of the Code in their mathematical difficulty, then the experiment would have little interest. For better or worse, however, the language featured in the two passages is standard Code drafting-speak, sprinkled throughout.

The Gain Question's statutory text, codified at 26 U.S.C. § 1001(a), contains the typical language the Code uses for subtraction.[70] As you can imagine, the Code does this "many times."[71] Moreover, § 1001(a) itself is one of the essential sections of the Code, defining "[t]he gain from the sale or other disposition of property,"[72] which is a fundamental concept of income tax.[73]

As for the language in the Annuity Question's statutory text, codified at 26 U.S.C. § 72(b), it so frequently bedevils tax-law readers

---

[69] Joseph Bishop-Henchman, *How Many Words Are in the Tax Code?*, TAX FOUND. (Apr. 15, 2014), https://taxfoundation.org/how-many-words-are-tax-code [https://perma.cc/3X74-GPUB] (estimating that 26 U.S.C. has 2,652 pages in their version, at around 450 words per page, i.e., 1,193,400 words).

[70] *See* GERSHON, *supra* note 8, at 10 ("Like addition, there are many times when the Code requires you to perform subtraction. The words that alert you that it is time to pull out your calculator and press the minus sign are 'the excess of . . . over.'" (omission in original)). One tricky thing about the language is that it is not *quite* subtraction: As the Gain Question exemplifies, if the subtraction would result in a negative number, the phrase instead picks out 0, rather than the negative number.

[71] *Id.* Indeed, my own search of a PDF of the Internal Revenue Code found 426 instances of the phrase "the excess of."

[72] 26 U.S.C. § 1001(a).

[73] For an admittedly rough indication of its essentiality, note that § 1001(a) is taught on page 2 of the standard Examples & Explanations hornbook. KATHERINE PRATT, THOMAS D. GRIFFITH & JOSEPH BANKMAN, FEDERAL INCOME TAX: EXAMPLES & EXPLANATIONS 2 (8th ed. 2019).

that one expert has given it its own name: "the infamous Code Ratio."[74] It features in Code sections as various as adoption expenses,[75] the Lifetime Learning credit,[76] the first-time homebuyer credit,[77] the renewable electricity production credit,[78] and the exclusion of gain from the sale of one's principal residence.[79]

With both of these constructions, and with other syntactic constructions denoting addition and multiplication,[80] the Code unfortunately overflows.

C. So What?

One could agree that using math symbols in the Code would make it easier to understand, without thinking that it matters. In this Section, I try to explain why it matters.

1. The Good Consequences of Understandability

There has been a recent concerted effort to make law easier to understand. Professors and organizations have advocated for a "plain English movement."[81] Committees have undertaken restyling projects

---

74 GERSHON, *supra* note 8, at 11. My search of a PDF of the Internal Revenue Code found 169 instances of the phrase "bears the same ratio."

75 26 U.S.C. § 23(b)(2)(A).

76 26 U.S.C. § 25A(d)(2).

77 26 U.S.C. § 36(b)(2)(A).

78 26 U.S.C. § 45(b)(1).

79 26 U.S.C. § 121(c)(1)(A).

80 For addition, the Code writes "the sum of . . . plus —." *See, e.g.*, 26 U.S.C. § 1001(b) ("The amount realized from the sale or other disposition of property shall be the sum of any money received plus the fair market value of the property (other than money) received."). For multiplication, the Code refreshingly just writes "multiplied by," though the surrounding language can make things a little complex. *See, e.g.*, 26 U.S.C. § 132(f)(6) ("[T]he dollar amounts . . . shall be increased by an amount equal to (i) such dollar amount, (ii) multiplied by the cost-of-living adjustment determined under section 1(f)(3) . . . ."). These constructions are intuitively less confusing than those for subtraction and division, though I still think that passages containing them would be more understandable if rephrased with '+,' '×,' and variables. If we wanted to get a more concrete quantitative estimate of the benefits of using math symbols in the Code, we could investigate these constructions as well, and construct a measure of how frequently sentences with these constructions are consulted when the Code is used.

81 *See, e.g.*, Carl Felsenfeld, *The Plain English Movement in the United States*, 6 CAN. BUS. L.J. 408, 408 (1981–82) (describing the early history of the movement); BRYAN A. GARNER, LEGAL WRITING IN PLAIN ENGLISH: A TEXT WITH EXERCISES 1 (3d ed. 2023) ("There's an age-old cycle of poor legal writing. You can help break it."); RICHARD C. WYDICK & AMY E. SLOAN, PLAIN ENGLISH FOR LAWYERS 5 (6th ed. 2019) (arguing that "good legal writing is plain English"); JOSEPH KIMBLE, LIFTING THE FOG OF LEGALESE: ESSAYS ON PLAIN LANGUAGE xv (2011) ("I think no reform would more fundamentally improve our profession and the work we do than learning to express ourselves in plain language. To that end, this book."); *About*, CLARITY INT'L, https://www.clarity-international.net/about.html [https://perma.cc/

to make various legal rules, such as the Rules of Evidence, more comprehensible.[82] And Congress has passed the Plain Writing Act, requiring various federal actors to explain things more clearly to the public.[83] One of the areas of law on which these advocates have focused, perhaps unsurprisingly, is the Code.[84]

The Internal Revenue Code affects an impressively large number of people. Every year, all U.S. citizens with income over certain amounts (e.g., $13,850, for a single 30-year-old in 2023) must file a tax return.[85] U.S. noncitizen residents with that much income must also file.[86] Similarly for corporations, partnerships, and various other organizations.[87] Foreign persons who come to the U.S. for part of the year for work must often file.[88] Foreign persons and organizations who do enough business in the U.S. must file.[89] And, since the U.S. implements much of its social

---

EG29-ML8L] (proclaiming themselves "a worldwide network of professionals who are committed to promoting plain legal language").

[82] *See, e.g.*, Sidney A. Fitzwater, *The Restyled Federal Rules of Evidence*, 53 WM. & MARY L. REV. 1435, 1438–42 (2012) (discussing the restyling of the Federal Rules of Evidence); Lisa Eichhorn, *Clarity and the Federal Rules of Civil Procedure: A Lesson from the Style Project*, 5 J. ASS'N OF LEGAL WRITING DIRS. 1 (3d prtg. 2008) (discussing the restyling of the Federal Rules of Civil Procedure).

[83] *See, e.g.*, *Law and Requirements*, PLAINLANGUAGE, https://www.plainlanguage.gov/law [https://perma.cc/UEM6-PCQB] (explaining the requirements of the Plain Writing Act).

[84] *See, e.g.*, Peter E. Boos, *Decoding the Code*, TAX NOTES 323, 323–24 (2017) (describing the complexity of the Code, and citing numerous calls for simplification).

[85] *Check if You Need to File a Tax Return*, IRS, https://www.irs.gov/individuals/check-if-you-need-to-file-a-tax-return [https://perma.cc/K46U-6F85] (listing the income thresholds for tax year 2023).

[86] *See, e.g.*, *U.S. Residents*, IRS, https://www.irs.gov/individuals/international-taxpayers/us-residents [https://perma.cc/9VKB-YUKD] (explaining that U.S. residents must generally obey the same tax rules as U.S. citizens); *Substantial Presence Test*, IRS, https://www.irs.gov/individuals/international-taxpayers/substantial-presence-test [https://perma.cc/56ND-2M79] (explaining one of the tests for who counts as a U.S. resident).

[87] *Business Taxes*, IRS, https://www.irs.gov/businesses/small-businesses-self-employed/business-taxes [https://perma.cc/2L6K-7V3N] ("All businesses except partnerships must file an annual income tax return. Partnerships file an information return.").

[88] *See, e.g.*, MINDY HERZFELD & RICHARD L. DOERNBERG, INTERNATIONAL TAXATION IN A NUTSHELL 37 (11th ed. 2018) (describing how "compensation for services performed in the United States is U.S. source income" which is generally taxable, with an exception for "compensation [that] does not exceed $3,000, a figure rendered virtually meaningless by the ravages of inflation since 1954, the year of enactment"). Some tax treaties between the U.S. and foreign states soften these requirements for taxpayer residents of the foreign states. *See id.* at 150–51 (explaining how tax treaties can ease the requirements for certain taxpayers).

[89] *See, e.g.*, *id.* at 22 ("[A] foreign corporation . . . is taxed on business profits from the conduct of a trade or business in the United States."); *id.* at 23 ("A nonresident alien individual or nonresident corporation that is a partner in either a U.S. or foreign partnership . . . is considered to be engaged in a trade or business within the United States if the partnership is so engaged."). An example is the notorious Revenue Ruling 58-63, which held that, if a foreigner permits a horse from their foreign racing stable to be raced in the United States, and the horse wins, this counts as being "'engaged in [a] trade or business within the United

assistance through the Tax Code, millions of citizens below the income threshold still file,[90] often paying a large chunk of their money to private businesses to assist them.[91]

Each year, the number of tax returns filed is around 250 million.[92] The total amount of time it takes to prepare all those returns is about 6.5 billion hours.[93]

Moreover, for better or worse, tax returns are not the only way we interact with the tax system. The Tax Code contains a number of taxes beyond the income tax.[94] And tax generates a large number of lawsuits,[95] which can usually be brought in front of not only specialized

---

States;'" consequently, "income derived therefore is subject to United States income tax unless exempted by some other provision of law." Rev. Rul. 58-63, 1958-1 C.B. 624.

[90] *See, e.g.*, Drew Desilver, *Who Pays, and Doesn't Pay, Federal Income Taxes in the U.S.?*, Pew Rsch. Ctr. (Apr. 18, 2023), https://www.pewresearch.org/short-reads/2023/04/18/who-pays-and-doesnt-pay-federal-income-taxes-in-the-us [https://perma.cc/9J6A-9ZC5] ("In 2020, the IRS received nearly 5.3 million individual tax returns that showed no AGI and hence no taxable income.").

[91] *See Why Do Low-Income Families Use Tax Preparers?*, Tax Pol'y Ctr., https://www.taxpolicycenter.org/briefing-book/why-do-low-income-families-use-tax-preparers [https://perma.cc/WMG6-DKSR] (Jan. 2024) ("Many low-income families owe no income tax but still must file a tax return to receive refundable tax credits . . . . Most [do so] with assistance from paid preparers . . . . That proportion is . . . 50.2 percent for returns with adjusted gross incomes below $30,000 . . . ."); Paul Weinstein Jr. & Bethany Patten, Progressive Pol'y Inst., The Price of Paying Taxes II: How Paid Tax Preparer Fees Are Diminishing the Earned Income Tax Credit (EITC) 1 (Apr. 2016), https://www.progressivepolicy.org/wp-content/uploads/2016/04/2016.04-Weinstein_Patten_The-Price-of-Paying-Takes-II.pdf [https://perma.cc/997N-CP3X] ("Workers eligible for the EITC continue to spend large sums—averaging around $400—at national tax preparation chains. In a recent survey . . . in Baltimore and Washington, D.C. we found that low-income taxpayers can expect to spend between 13 and 22 percent of the average EITC refund to file their taxes."). *See also* Stacy Cowley, *Tax Refund Loans Are Revamped and Resurrected*, N.Y. Times (Jan. 15, 2017), https://www.nytimes.com/2017/01/15/business/tax-refund-loans-are-revamped-and-resurrected.html [https://perma.cc/8RXT-R8FB] ("Tax preparation prices are frequently opaque, and wildly variable . . . . 'Only one office even had prices disclosed publicly, and often, the fees seemed to be completely arbitrary,' said Liz Coyle, Georgia Watch's executive director. 'It's almost impossible for someone to comparison-shop.'").

[92] *Returns Filed, Taxes Collected & Refunds Issued* (Apr. 18, 2024), https://www.irs.gov/statistics/returns-filed-taxes-collected-and-refunds-issued [https://perma.cc/RHU8-ETGB] (displaying the number of different types of returns filed in fiscal year 2023).

[93] Dan Bosch, *Tracker: The Cost of Tax Paperwork*, Am. Action F. (Apr. 14, 2022), https://www.americanactionforum.org/insight/tracker-the-cost-of-tax-paperwork [https://perma.cc/UW4L-6BK4].

[94] *Title 26, U.S. Code*, U.S. Census Bureau, https://www.census.gov/history/www/reference/privacy_confidentiality/title_26_us_code_1.html [https://perma.cc/WSW5-J8N9] ("The Internal Revenue Code (IRC) is the body of law that codifies all federal tax laws, including income, estate, gift, excise, alcohol, tobacco, and employment taxes.").

[95] Precise numbers are hard to come by, but Casetext has about 75,000 Tax Court cases from 2023. *See Tax Court Cases by Year*, Casetext, https://casetext.com/cases/ustc [https://perma.cc/EH8Y-TDBW] (totaling over 75,000 cases in 2023 and over 84,000 cases in 2022). A very rough estimate is that 95% of tax cases are filed in Tax Court. *See* L. Paige Marvel, *The*

judges on the Tax Court, but also generalist judges and juries in federal district courts.[96] These non-tax-expert judges, their clerks, jurors, and appellate judges must work through the dense Code, understandably but unfortunately making mistakes along the way.[97]

All this behavior revolves around the Tax Code at the center. It is the Code which dictates tax liabilities, serves as the foundation of an enormous mass of regulations and informal documents, gets fought over by lobbyists and politicians, and is interpreted by all the advisors, lawyers, and judges who come within its reach.[98]

Unfortunately, as the examples in Part II illustrate, our tax system is hard to understand.[99] This difficulty creates negative effects. Take, for example, the implementation of the Earned Income Tax Credit. Passed under President Ford in 1975,[100] and significantly expanded since,[101] the credit is used by the federal government to distribute a total of around $60 billion each year to 25 million lower-income families.[102] This is one

---

Evolution of Trial Practice in the United States Tax Court, 68 Tax Law. 289 (2015) (stating that "the Tax Court adjudicates more than 95% of the tax cases filed by taxpayers nationally").

96 Joseph A. Bankman, Daniel N. Shaviro, Kirk J. Stark & Edward D. Kleinbard, Federal Income Taxation 41 (18th ed. 2019). For example, before it was decided by the Supreme Court, the tax case *Moore v. United States*, was brought in federal district court before a generalist judge, and then appealed to the Ninth Circuit, again before judges who would not normally have a tax background. Moore v. United States, 36 F.4th 930, 934 (9th Cir. 2022) (stating the procedural background of the case). Cases can also be brought in the United States Court of Federal Claims. Bankman et al., *supra* at 96.

97 *See, e.g.*, Boos, *supra* note 84, at 323 & n.3 (citing cases where circuit courts clearly misapplied the law).

98 *See, e.g.*, Bankman et al., *supra* note 96, at 42–45 (describing the sources of federal tax law).

99 *See, e.g.*, Boos, *supra* note 84, at 324. And people don't like the complexity. Desilver, *supra* note 90 ("In a recent Pew Research Center survey, 53% of U.S. adults said the system's complexity bothered them a lot.").

100 Cong. Rsch. Serv., The Earned Income Tax Credit (EITC): A Brief Legislative History, at ii (2018), https://crsreports.congress.gov/product/pdf/R/R44825/8 [https://perma.cc/EE5N-6VRF].

101 *Id.* ("After various legislative changes over the past 40 years, the credit is now one of the federal government's largest antipoverty programs.").

102 *See Statistics for Tax Returns with the Earned Income Tax Credit (EITC)*, IRS, https://www.eitc.irs.gov/eitc-central/statistics-for-tax-returns-with-eitc/statistics-for-tax-returns-with-the-earned-income [https://perma.cc/Q5UQ-CF5Y] (stating that, for the tax years 2022, '21, and '20, the amount of claims were $57, 64, and 60 billion, and the number of claims were 23, 31, and 25 million workers and families, respectively). The credit is a "refundable" credit, which means that it not only reduces the amount of taxes you have to pay (with $1 of tax credit getting "used up" to erase $1 of tax liability), but also any remaining amount of the credit gets refunded to you in cash if you have no more tax to pay. *See generally What is the Difference Between Refundable and Nonrefundable Credits?*, Tax Pol'y Ctr., https://www.taxpolicycenter.org/briefing-book/what-difference-between-refundable-and-nonrefundable-credits [https://perma.cc/3H7W-69KN] (answering the title's question).

of the main ways our government provides a social minimum: a floor under which we believe it is unjust to allow our fellow compatriots to fall.[103] The amount a person receives depends on their marital status, how many children they have, and how much they earn (with the amount growing, plateauing, and then shrinking, as their earnings rise).[104] To get a feel for the numbers, for tax year 2023, a single person who earns $10,000 and has no children is entitled to $583, and a married family who earns $20,000 and has two children is entitled to $6,604.[105] The money is not distributed automatically, however—individuals are required to file a tax return in order to receive the credit.[106] Each year, about 5 million eligible people fail to claim the credit, so around $7 billion of eligible funds are never disbursed.[107] Those who do file a tax return often pay about $400 of the credit to businesses who fill out the returns for them.[108] Yes, making the Code easier to understand would still leave many of these problems unsolved, but it would be a good start, with possible ripple effects on the generalist judges, regulations, government websites, and advisors who all try to understand the Code and explain it to others.[109]

Furthermore, the current complexity itself creates a perverse feedback loop. The tax system's difficulty increases demand for tax preparers, who then have a financial incentive to keep the complexity in place or even increase it.[110] One notorious example of this was the

---

103 *See What is the Earned Income Tax Credit?*, Tax Pol'y Ctr., https://www.taxpolicycenter.org/briefing-book/what-earned-income-tax-credit [https://perma.cc/9EMB-53KZ] ("[I]f the EITC were treated like earnings, it would have been the single most effective antipoverty program for working-age people, lifting about 5.6 million people out of poverty in 2018, including 3 million children . . . ."). *See generally* Stuart White, *Social Minimum*, *in* Stanford Encyclopedia of Phil. (Winter 2021 ed.), https://plato.stanford.edu/entries/social-minimum [https://perma.cc/CM7G-DU82] (explaining the concept of a social minimum, and evaluating arguments for and against it). Of course, we have arguably set the floor far too low.

104 *Policy Basics: The Earned Income Tax Credit*, Ctr. on Budget & Pol'y Priorities (Apr. 28, 2023), https://www.cbpp.org/research/policy-basics-the-earned-income-tax-credit [https://perma.cc/EK9F-X4F6].

105 *Id.* (entering the pertinent values into the figure entitled "Value of Federal Earned Income Tax Credit, 2023").

106 *Do All People Eligible for the EITC Participate?*, Tax Pol'y Ctr., https://www.taxpolicycenter.org/briefing-book/do-all-people-eligible-eitc-participate [https://perma.cc/EML5-XS66].

107 *Id.*

108 *See* Weinstein & Patten, *supra* note 91.

109 An example of some of the math in the Earned Income Tax Credit section of the Code: "[the allowable credit] for any taxable year shall not exceed the excess (if any) of . . . the credit percentage of the earned income amount, over . . . the phaseout percentage of . . . the adjusted gross income . . . of the taxpayer for the taxable year as exceeds the phaseout amount." 26 U.S.C. § 32(a)(2).

110 For an expression of this idea, see, for example, Len Burman, *The Tax Complexity Lobby*, Forbes (Apr. 15, 2013, 3:05 PM), https://www.forbes.com/sites/leonardburman/2013/04/15/

battle for a return-free tax system. In other high-income countries, such as the United Kingdom, the government calculates most people's taxes for them, sends them the details, and then the individual taxpayers are free to accept or amend these pre-filled-out returns.[111] The Obama Administration tried to change the United States to such a system, but Intuit (the maker of TurboTax) and H&R Block, the two dominant tax-preparation software firms, spent millions to lobby against such a change, and they won.[112] The complexity continues.[113]

If we can reduce this complexity, even in just one aspect, the positive effects could multiply, because the Code affects so many people, so often. And we could add elementary math symbols not only to the laws themselves, but also to regulations and informal documents supported by those laws and published by the Department of Treasury and the Internal Revenue Service.[114]

---

the-tax-complexity-lobby [https://perma.cc/CF6Z-H38R] (positing that tax preparers would lose profit if filers could do their returns without help).

111 *See* Thomas J. Healey, *Return-Free Tax Filing: A Good Idea Whose Time Has Come and Come and Come*, Milken Inst. Rev. (Dec. 15, 2022), https://www.milkenreview.org/articles/return-free-tax-filing [https://perma.cc/8J27-ZM8D] ("Indeed, an estimated 36 countries—including the United Kingdom, Germany, Japan, Spain, Denmark and Sweden—have implemented return-free filing.").

112 *See* Dylan Matthews, *Why I'm Boycotting TurboTax this Year*, Vox (Apr. 15, 2019, 1:58 PM), https://www.vox.com/2016/3/29/11320386/turbotax-boycott-lobbying-tax-filing-season-tax-day-april-15 [https://perma.cc/5439-J4C8] ("Years ago, the Obama administration proposed a system of automatic tax filing, in which the IRS uses income information it already has to fill out your tax return for you . . . . [T]he idea has gone nowhere. The main reason? Lobbying from Intuit and H&R Block."). *See generally* Justin Elliott & Paul Kiel, *Inside TurboTax's 20-Year Fight to Stop Americans from Filing Their Taxes for Free*, ProPublica (Oct. 17, 2019, 5:00 AM), https://www.propublica.org/article/inside-turbotax-20-year-fight-to-stop-americans-from-filing-their-taxes-for-free [https://perma.cc/W8F9-PY7J] (detailing the long history of Intuit's successful lobbying to stop Congress from simplifying the tax system); Jessica Huseman, *Filing Taxes Could Be Free and Simple. But H&R Block and Intuit Are Still Lobbying Against It*, ProPublica (Mar. 20, 2017, 1:22 PM), https://www.propublica.org/article/filing-taxes-could-be-free-simple-hr-block-intuit-lobbying-against-it [https://perma.cc/M5SR-4R78] (describing both H&R Block and Intuit's continued lobbying to stop the IRS from pre-filling out tax returns). President Reagan first suggested a return-free system back in 1985. Beverly Moran, *Why Can't the IRS Just Send Americans a Refund – or a Bill?*, The Conversation (Mar. 22, 2021, 8:27 AM), https://theconversation.com/why-cant-the-irs-just-send-americans-a-refund-or-a-bill-156733 [https://perma.cc/24P4-SESF].

113 *But see* Ann Carrns, *I.R.S. to Begin Trial of Its Own Free Tax-Filing System*, N.Y. Times (Jan. 5, 2024), https://www.nytimes.com/2024/01/05/your-money/irs-tax-filing-free-online.html [https://perma.cc/EDM7-TLX6] (describing how the IRS is letting certain taxpayers in twelve states try a new, free online filing method, amidst opposition from Republicans and tax-preparation companies).

114 In contrast to the Code, the regulations do occasionally use math symbols, especially when "showing their work" in illustrative examples. But they almost always describe the rules themselves using solely English words. *See, e.g.*, 26 C.F.R. §§ 1.0–1.60, https://www.govinfo.gov/content/pkg/CFR-2023-title26-vol1/PDF/CFR-2023-title26-vol1.pdf [https://perma.cc/P9NA-KCEV] (where a pdf search for '÷' yields 12 matches in the 743 pages). For example,

2. The Democratic Value of Understandability

Finally, it is important to recognize that, even apart from beneficial effects that might be measured in a cost-benefit analysis, there is further value in making our laws more understandable.

It is a fundamental principle of the rule of law and the democratic spirit that, all else equal, people should be able to understand the laws that govern them. John Rawls, for example, argued that the rule of law requires "public rules addressed to rational persons" and that "laws be known and expressly promulgated, that their meaning be clearly defined."[115] Lon Fuller similarly argued that the rule of law requires "legislative clarity" as opposed to "obscurity."[116] And Chief Justice Marshall wrote that "[t]he rules by which the citizen shall be deprived of his liberty or property, to enforce a judicial sentence, ought to be prescribed and known . . . ."[117]

Some laws, such as regulations concerning nuclear power plants, may have to be difficult to understand. But surely one can agree that, all else equal, it is a virtue of a law in a democracy if it is comprehensible to the public. Yes, this may sometimes have to be sacrificed to achieve other virtues, such as efficacy. Really, it is hard to see how effective nuclear power plant regulations could be written in such a way that an ordinary person could, with very little additional work, understand them. But if one does not have to sacrifice the virtue of general comprehensibility, one should not.

---

one regulation first spells out a general rule as follows: "The child's share of the allocable parental tax is an amount that bears the same ratio to the total allocable parental tax as the child's net unearned income bears to the total net unearned income of all children of such parent to whom section 1(i) applies." Treas. Reg. § 1.1(i)-1T (1987). It then gives a concrete example, again using words to describe the general rule, but using symbols at the end to show how it got the answer: "*Example 2.* H and W have 3 children, A, B, and C . . . . Each child's share of the allocable parental tax is an amount that bears the same ratio to the total allocable parental tax as the child's net unearned income bears to the total net unearned income of A, B, and C. Thus, A's share . . . is $1,650 (5,000 ÷ 10,000 × 3,300) . . . ." *Id.* IRS publications similarly use English words when describing rules. *See, e.g.*, IRS, Pub. No. 575, Pension and Annuity Income 14 (2022), https://www.irs.gov/pub/irs-pdf/p575.pdf [https://perma.cc/6EX5-PJVL] ("Under the General Rule, you determine the tax-free part of each annuity payment based on the ratio of the cost of the contract to the total expected return."). However, when discussing examples they will often use math symbols, or English closer to math, or they will simply direct you to "worksheets" where you are instructed to write down numbers and then do some arithmetic. *See, e.g.*, *id.* ("Bill's tax-free monthly amount is $100 ($31,000 ÷ 310) as shown on line 4 of the worksheet.").

115 John Rawls, A Theory of Justice 207–09 (rev. ed. 1999).

116 Lon L. Fuller, The Morality of Law 63–65 (rev. ed. 1969). "Clarity" can connote both precision and ease of understanding, but both aspects are important, because the ultimate aim is for understanding, and there will be little collective understanding if a law is, for example, very precise but so syntactically complex that only one in a billion can parse it.

117 Wayman v. Southard, 23 U.S. (10 Wheat.) 1, 13 (1825).

The key point here is that using elementary math symbols would not significantly change the *content* of the Code, but merely its *presentation*. I am not advocating for the use of significantly different concepts, as would be required if we, say, got rid of the four technical mens rea mental states in the Model Penal Code,[118] or the scientific concepts used in the regulations of nuclear power plants. I am instead advocating that we refer, for example, to division not by the symbol 'ratio,' but rather by the symbol '÷.' A closer analogy to my proposal would be advocating for the replacement of a complicated Gothic font by a more modern, readable one, or advocating for the internet publication of regulations that were previously print-only.[119] If we *can* make laws more accessible, without significantly changing them, then, especially in a democracy and for laws concerning almost everyone in the country, such as the Code, we ought to do so.

One might object that it is perfectly fine if laws are addressed to experts, as long as they then explain their approximate meaning in simpler terms to the masses. This is arguably the current spirit of much of the Code. In contrast, I would propose that this violates the spirit of democracy. The general laws of a truly just society should be understandable by almost all people in that society. It may be true that this far-reaching claim is not necessary here, where there is little to be gained by avoiding the use of elementary mathematical symbols in the Code. But it is an admirable vision that could animate efforts to make the United States Code in general more understandable to all.[120]

3. Answers to Additional Objections

One might object that few people actually look at the Code, so there is little to be gained by making it more understandable.[121] While it's hard to find evidence about the number of people who read U.S.

---

118 MODEL PENAL CODE § 2.02(2) (AM. LAW INST., Proposed Official Draft 1962) (listing purpose, knowledge, recklessness, and negligence as the four mental states of mens rea).

119 An even closer analogy is the restyling of the Federal Rules of Evidence, which was intended "to make the Rules simpler, easier to read, and easier to understand without changing their substance." Fitzwater, *supra* note 82, at 1440.

120 In *Animal Farm*, many of the animals can't read the laws, so the pigs interpret the laws for them. It doesn't end well. GEORGE ORWELL, ANIMAL FARM: A FAIRY STORY 49 (1945) ("None of the other animals on the farm could get further than the letter A . . . . [T]he stupider animals . . . were unable to learn the Seven Commandments by heart . . . . Snowball declared that the Seven Commandments could . . . be reduced to a single maxim . . . 'Four legs good, two legs bad.'").

121 For a distinct but related objection, see Samuel A. Donaldson, *The Easy Case Against Tax Simplification*, 22 VA. TAX REV. 645, 692 (2003) (claiming that it is "idealistic to believe that taxpayers as a whole understand" or "even care about" central issues of Federal tax policy).

statutes, our everyday experience with nonlawyers suggests that the number is indeed low. But there are still benefits to increasing its understandability.

First, as mentioned earlier,[122] some nonexperts must look at the Code: for example, pro se taxpayers, generalist judges, their clerks, and jury members. This Note's experiment provides evidence that using math symbols in the Code would help them perform their task. And given the tens of thousands of lawsuits every year,[123] this could mean helping tens of thousands of *people* every year.

Second, this may well be idealistic, but if our laws were clearer, perhaps more people would look at them.[124] So it seems a bit perverse to say that we needn't make our laws understandable because people don't look at them.

Third, and most importantly, this objection forgets the value of having a more democratic and rule-of-law-bound society through the promulgation of laws that are understandable.[125] Our ideal vision of society is not of some Platonic oligarchy, where we trust the philosopher-rulers to pass the right laws and tell us what they mean. Rather, we believe that, if the law states that you must do something, it ought to do so in a way you can understand.

One might also object that the proposal would take too much expensive lawyer time to implement.[126] I think there's some truth here. Most individual passages could be translated fairly easily, especially since tax lawyers usually perform the translation in their heads anyway,

---

122 *See supra* notes 96–97.

123 *See supra* note 95.

124 Perhaps understandably, no one seems to have performed a quantitative study about the percentage of nonlawyers who read statutes, or the causes for the particular percentage being what it is. So again we must rely here on our common sense and anecdotal experience. And anecdotally, people do state that one of the reasons for not looking at the laws is their lack of clarity. *See, e.g.*, Erin Dixon, *Language of Law Can Be Confusing to Many*, MIDVALE J. (Mar. 29, 2022, 9:45 PM), https://www.midvalejournal.com/2022/03/29/394022/language-of-law-can-be-confusing-to-many [https://perma.cc/9HFS-Z5C8] ("'The legal jargon [in bills] is so overwhelming sometimes that it makes me want to give up reading it entirely,' Midvale resident KoriAnne Starr Lucero said. 'How are we as constituents supposed to know what is being passed . . . if it's intentionally made difficult to understand?'"). In econ-speak, "demand curves slope downward": Usually, if something is more costly, including in time or mental effort, people choose less of it. *See generally* Samuel Bowles & Sandra Polanía-Reyes, *Economic Incentives and Social Preferences: Substitutes or Complements?*, 50 J. ECON. LIT. 368, 369–70 (June 2012 ed.) (discussing empirical evidence of when and how incentives affect behavior).

125 *See infra* Section II.C.2.

126 *See* Boos, *supra* note 84, at 337–38 (discussing the objection that tax simplification in general would be too costly).

as a first step in understanding a particular section.[127] But to do this for the approximately 1.2 million words in the Code as a whole would be an unenviable project.[128] One alternative is that we could decide just to use math symbols going forward in new legislation. Old laws could stay as they are, but new tax bills—and tax is one of the fastest changing areas of the law[129]—could help themselves to elementary math symbols. Intuitively, drafting a passage using math symbols would perhaps be even easier than drafting one using just English, and, since future lawyers would find it easier to read, the net result would actually be a saving of legal effort.

Finally, one might worry that the proposal could open the door to more complicated math symbols that would actually *increase* the complexity and difficulty of the Code, such as using the number *e* in calculating continuous compound interest. I do think this would be troubling, which is why I have stressed using only elementary math symbols, focusing on the basic signs '+', '–', '×', and '÷'. Consequently, if this proposal were to be taken up, the emphasis should be on the gains in understandability, including the democratic value of clearer laws. Such an emphasis would hopefully act as an antidote to any attempt to use symbols to make the code more complex.

Conclusion

Our tax code instructs us to add, subtract, multiply, and divide. But it surprisingly does so using tortuous, hard-to-apply English phrases instead of the signs taught in elementary school. This Note's preregistered, randomized, controlled trial involving 161 participants found that using elementary math symbols in actual tax statutes substantially increased problem-solving accuracy: raising it from 25% to 70% in one question, and from 11% to 50% in another. Using these symbols in our tax statutes would have beneficial effects and would further the rule of law and the spirit of democracy. So we should do it.

---

127 See, for example, *supra* note 114, where a regulation translates the English into math symbols in order to show the correct answer to a problem.

128 *See* Bishop-Henchman, *supra* note 69.

129 *See* Rebecca M. Kysar, *Reconciling Congress to Tax Reform*, 88 Notre Dame L. Rev. 2121, 2123 (2013) ("Tax law . . . constantly churns, somehow avoiding the molasses of the legislative process. A common critique levied against tax law is that there is *too much* legislative action, resulting in ever-changing rules."). One of the main reasons tax law changes so frequently is because tax bills, unlike many other bills, can be passed in the Senate by a simple majority through the reconciliation process. *See What is Reconciliation?*, Tax Pol'y Ctr., https://www.taxpolicycenter.org/briefing-book/what-reconciliation [https://perma.cc/R69J-2JJ4] (describing reconciliation and how it was used to pass various landmark tax bills).