Misconceptions about the long division algorithm in school mathematics*

Hung-Hsi Wu

January 25, 2021

This article has appeared in *Journal of Mathematics Education at Teachers College,* Volume 11, Issue 2, Fall 2020, 1-12. `https://jmetc.columbia.edu`

For roughly the last five decades, our nation has had a *de facto* national school mathematics curriculum, one that has been defined by the standard school mathematics textbooks.[1] It is a fact, though not one that has been explicitly discussed in the world of mathematics education, that the mathematics embedded in these textbooks is extremely flawed, to the point of being unlearnable.[2] It is notable for its lack of definitions for concepts (e.g., what is a fraction and what does it mean to multiply two fractions?), lack of reasoning for skills (e.g., why is negative times negative positive?), almost universal lack of precision (e.g., is "$3^0 = 1$" a definition or a theorem?), a general lack of coherence (e.g., are finite decimals and fractions different kinds of numbers?), and a pervasive lack mathematical purpose in its presentation (e.g., telling students to learn to take the absolute value of a number by killing the negative sign because this skill will be on the test). For ease of exposition, we call this particular version of school mathematics **TSM** (Textbook School Mathematics). We will refer to pp. 22-30 (*of the pagination of the PDF*) of Wu, 2020d for a more detailed discussion of TSM. In retrospect, much of the turmoil in school mathematics education during the last thirty years has revolved around disagreements on how best to deal with the absurd situations that arise in our school

*I am very grateful to Larry Francis for his editorial assistance.

[1]This statement does not take into account a very small number of textbooks in the last five years that *may be* exceptions.

[2]We use "unlearnable" in this article to mean "unlearnable by a majority of students", and *learning mathematics* is understood to include learning how to reason.

mathematics classrooms when teachers try to teach something as nonsensical as TSM (see Wu, 2020d). For all these reasons (and more), it is no longer possible in 2020 to discuss mathematics learning in schools without directly confronting TSM. The central issue now is how to eradicate TSM and help teachers and students, schools and districts transition to a different version of K-12 mathematics that is transparent, and therefore learnable.

There is no better illustration of the fiasco that is TSM than the multiple defects in how the concept of division is taught in elementary school. The ubiquity of the limerick, "Ours is not to reason why, just invert and multiply", points to the catastrophic failure in the teaching of the *division of fractions*, but less well-known is the fact that a failure of comparable magnitude has already occurred in the teaching of division among the whole numbers. On the one hand, there is the concept of *the division of one whole number by another,* such as $35 \div 7$ or $36 \div 6$, and on the other, there is the concept of the *division-with-remainder of one whole number by another*, e.g., the division-with-remainder of 35 by 6 for which the symbol $35 \div 6$ cannot be used. These are two different concepts but TSM makes believe that they can be conflated. In addition, TSM does not make explicit the fact that the long division algorithm, e.g., of 35 by 6, is a shortcut that yields the division-with-remainder of 35 by 6. These flaws of TSM destroy the bridge that leads from the long division algorithm to at least two topics in middle and high school: the conversion of a fraction to a decimal by "the long division of the numerator by the denominator"[3] and the division algorithms for polynomials. This is but one example of how TSM suppresses the *coherence* of school mathematics and, instead, presents mathematics to students as a collection of fragments and factoids to be memorized by brute force.

The mishandling of the long division algorithm by TSM in elementary school and the ripple effects of this particular failure in the school mathematics curriculum are the main concerns of this article. It will also offer some suggestions on how to improve the teaching of this algorithm in grades 4-6. In the last section, we put this discussion of the long division algorithm in the broader context of how TSM has made school mathematics a horror story. We will also describe a recent development—the publication of a detailed curricular road map for making K-12 mathematics *mathematical*—that may eventually render TSM a relic of the past.

[3]The phrase in quotes requires a detailed explanation which will be given later.

Teaching long division in grade 4

Consider teaching the long division 78 by 4 in grade 4. The usual setup for long division is to draw a "division house" (in the terminology of Green, 2014), putting 78 inside and 4 outside.

$$\begin{array}{r r r} & 1 & 9 \\ \cline{2-3} 4\ \big) & 7 & 8 \\ & 4 & \\ \cline{2-3} & 3 & 8 \\ & 3 & 6 \\ \cline{2-3} & & 2 \end{array} \tag{1}$$

Students are told that the number 19 on the roof and the 2 at the bottom are the answer to the following question: if they want to put 78 apples in groups of 4, how many such groups are there and how many apples (if any) are left over? They are also taught to write this as $78 \div 4 = 19\,R2$.

Fourth graders undoubtedly have a hard time understanding why the "division house" in (1) gives the correct answer of 19 equal groups of 4 with remainder 2. The effect of the putative "equality" $78 \div 4 = 19\,R2$ on their mathematics learning is, however, more insidious and more lasting. First of all, mathematics education certainly should not engage in teaching something that is blatantly false, but $78 \div 4 = 19\,R2$ *is* blatantly false. To see this, if we divide 59 by 3, we also get 19 with remainder 2. So we have $59 \div 3 = 19\,R2$. It follows that $59 \div 3$ and $78 \div 4$ (whatever they are!) must be equal since they are both equal to $19\,R2$. Even fourth graders can sense that, whatever "equality" means, the equality $59 \div 3 = 78 \div 4$ looks really bad. To understand why, we have to put ourselves in the context of fractions to see that this implies $\frac{59}{3} = \frac{78}{4}$, which implies $19\frac{2}{3} = 19\frac{2}{4}$, which in turn implies $\frac{2}{3} = \frac{2}{4}$. The last equality is clearly false.

Let us look at $78 \div 4 = 19\,R2$ from a different angle. It would appear that TSM uses it as a shorthand for "do the long division of 78 by 4 and the answer is 19 with remainder 2". It turns out that this kind of illegitimate shorthand is part of a common pattern in TSM. Consider the teaching of fractions in TSM, for example. The equal sign in any of the formulas for arithmetic operation is always used as a call to do a computation or announce the result of a computation (in the following, a, b, etc., are whole numbers which may be assumed to be nonzero where necessary):

$$\frac{a}{b} \pm \frac{c}{d} = \frac{pa \pm qc}{m} \quad \text{where } m \text{ is the least common multiple of } b \text{ and } d \text{ and } m = pb = qd \tag{2}$$

$$\frac{a}{b} \times \frac{c}{d} = \frac{ac}{bd} \tag{3}$$

$$\frac{a/b}{c/d} = \frac{a}{b} \times \frac{d}{c} \tag{4}$$

None of these formulas is intended by TSM to convey the message that the quantities on the two sides are "equal". Indeed, a correct mathematical exposition would first *define* what it means to add, subtract, multiply, and divide two fractions before *proving* that the fractions on both sides of each of (2)–(4) are the same fraction, i.e., the same point on the number line. (See Wu, 2011, Sections 14.1, 16.1, 17.1, and 18.2.) However, TSM as a rule does not provide *definitions* for concepts—it may provide pictures and metaphors, but not mathematical definitions—so that students are left in the dark about what a fraction is and, therefore, also what it means to add, subtract, multiply, and divide two fractions. In TSM, each of these four equations, like $78 \div 4 = 19\,R2$, is nothing more than a command to perform a computation, e.g., (4) says to divide $\frac{a}{b}$ by $\frac{c}{d}$, simply invert $\frac{c}{d}$ and multiply by $\frac{a}{b}$. In TSM, it is irrelevant what "division" means; all that matters is that students get the right answer when called upon to do a division.

To further reinforce our claim that TSM consistently misinforms students about the equal sign, let us look at what TSM says about equations and how to solve them. According to TSM, an equation (in one variable) is an equality of two expressions involving a "variable" x, such as $3x + 1 = x - 5$. The instruction from TSM on how to solve such an equation is to go through the following steps of symbolic manipulations:

Step A: $(-x) + 3x + 1 \;=\; (-x) + x - 5$

Step B: $2x + 1 \;=\; -5$

Step C: $2x + 1 + (-1) \;=\; -5 + (-1)$

Step D: $2x \;=\; -6$

Step E: $x \;=\; -3$

The answer of -3 is indeed correct, but what do steps A–D *mean*? Take Step A, for example. TSM says it follows from the *equality* $3x + 1 = x - 5$ by adding the same expression $(-x)$ to both sides. But in what sense is $3x + 1 = x - 5$ an equality? Since x is a variable, it can take on arbitrary values such as $x = 1$. In that case, the left side is

4 and the right side is -4, and they are certainly not equal! The same comment applies to Steps B, C, and D. The use of the equal sign in this standard process of solving the linear equation is therefore a *mathematical* travesty. So, once again, what TSM wants is not for students to learn how to correctly use the equal sign but only to know that they should go on automatic pilot to do computations at the sight of the equal sign. (For a correct definition of what an equation is and how to correctly solve the equation $3x + 1 = x - 5$ via Steps A–D, see Sections 2.1 and 3.1 of Wu, 2016b.)

There should thus be no mystery about why students fail to understand the meaning of the equal sign: TSM *has systematically corrupted their conception from the outset. Garbage in, garbage out.* This failure has drawn the attention of educators in the past four decades because it has hampered students' ability to learn algebra in middle school (e.g., Falkner et al., 1999, Kieran, 1981, and Knuth et al., 2008). However, the connection between TSM and students' failures in mathematics, particularly algebra, seems to have been overlooked thus far. As mentioned above, one cannot look past TSM in year 2020 in any attempt to improve student learning, so we hope education research will at last recognize the need to eradicate TSM from school mathematics education.

Let us now revisit (1). A key point is how to correctly introduce the division symbol "$\div$" to students *in the context of whole numbers.* We define $35 \div 5$ to be the whole number k so that $35 = k \times 5$ (*in exactly the same way* that we introduce the subtraction $17-9$ to be the whole number $\square$ so that $9+\square = 17$).[4] Then it is clear that the equation $35 \div 5 = 7$ is correct since $35 = 7 \times 5$. In general, if we know ahead of time that m is a multiple of n ($n \neq 0$), then $m \div n$ is *by definition* the whole number k so that $m = k \times n$. If, however, m is *not* multiple of n, we are at a loss as to what $m \div n$ could mean as a whole number or two whole numbers. There is therefore no way that something like $78 \div 4 = 19\,R2$ could make any sense as an equality about whole numbers—unless you *insist*, as TSM does, that the *computation with the "division house" must have an answer* and "$78 \div 4 = 19\,R2$" is the symbolic expression of choice. School mathematics must reject such bizzare impulses and teach students to rigorously observe the fact that—*in the context of whole numbers*—the division symbol $\boldsymbol{m \div n}$ ($n \neq 0$) can be used *only* when m is known to be a multiple of n. This kind of precision is by no means inappropriate for fourth graders. After all, even second graders learn not to write $5 - 9$, or in general

[4]Informally, we sometimes rephrase this statement about division as follows: $35 \div 5 = 7$ as an alternate way of expressing $35 = 7 \times 5$.

$k - \ell$ when $k < \ell$, in the context of whole numbers.

If we do not know whether or not m is a multiple of n, then we have to introduce the concept of *division-with-remainder*. Here is the definition of the **division-with-remainder of m by n** ($n \neq 0$): it is an expression of m in terms of n and two whole numbers q and r so that

$$m \;=\; (q \times n) + r \quad \text{where } 0 \leq r < n \tag{5}$$

The number q is called the **quotient** of the division-with-remainder and r its **remainder**.[5] If it happens that the remainder r in the division-with-remainder is 0, then m is a multiple of n and the two concepts of *division of m by n* and *division-with-remainder of m by n* coincide. We note that the restriction of $0 \leq r < n$ on the remainder r is an essential part of the definition because it guarantees that the whole number q is the *largest* whole number so that $q \times n \leq m$, as we now explain.

In fourth grade, of course we give the definition of division-with-remainder only by the use of explicit examples. For the case at hand: the division-with-remainder of 78 by 4 is expressed as

$$78 \;=\; (19 \times 4) + 2 \quad \text{where the "2" satisfies } 0 \leq 2 < 4 \tag{6}$$

This equation clearly implies that if there are 78 apples (i.e., the left side of (6)), then it is the same number of apples as in 19 groups of 4 apples (i.e., (19×4)) plus 2 extra apples on the side (i.e., the $+2$ on the right side of (6)). *This is the intuitive meaning of "division-with-remainder of 78 by 4" that we want to convey to students.* If we can teach students how the "division house" in (1) leads directly to (6), then the "division house" will become learnable mathematics rather than just a senseless ritual.

Naturally, there are other expressions for 78 that superficially resemble (6). For example,

$$78 \;=\; (18 \times 4) + 6$$

But this is not the division-with-remainder of 78 by 4 because the "remainder" here, 6, does not satisfy the requirement of being < 4 as stipulated in (6). So we take out another "group of 4 apples" among the 6 left over so that the 18 equal groups of 4 become 19 equal groups, and there are now 2 left over as in (6). On the other hand, we cannot get

[5] For lack of space, we will skirt the issue of the *uniqueness* of both q and r; this is discussed, e.g., on pp. 104-105 of Wu, 2011.

20 equal groups of 4 out of 78 because $20 \times 4 = 80$, which is greater than 78. Therefore the 19 in (6) is the largest whole number so that $(19 \times 4) \leq 78$. We usually express this by saying that **19 is the largest multiple of 4 that is $\leq$ 78**.

By tradition, we continue to call m the **dividend** and n the **divisor** in (5). Thus, 78 is the dividend, 4 is the divisor, 19 is the quotient, and 2 is the remainder in the division-with-remainder (6).

Knowing that the quotient is just the largest multiple of the divisor not exceeding the dividend tells us that no thinking is needed to get the division-with-remainder of one number by another. For example, to find the division-with-remainder of 78 by 4, we could simply write out the multiples of 4 until we get close to 78:

$$0, 4, 8, 12, \ldots, 68, 72, 76, 80, \ldots$$

By inspection, 76 is *that* multiple. So, since $76 = 19 \times 4$ and $78 - 76 = 2$, we see that the division-with-remainder of 78 by 4 is given by (6), i.e.,

$$78 \;=\; (19 \times 4) + 2$$

While this way of getting the quotient and remainder may be straightforward, it can get very tiresome very fast: think about getting the division-with-remainder of 78765 by 4 by listing all the multiples of 4 up to and just beyond 78765. We need a shortcut, and the "division house", i.e., the long division algorithm in (1) is that shortcut, as we now show. After all, (1) is a bit more pleasant than listing the multiples of 4 up to 80.

A main purpose of this article is to tell the full story about *why* the long division algorithm in (1) leads inexorably to (6), but grade 4 may not the right place to do this. Nevertheless, if we believe in teaching students *mathematics* rather than just procedures, we have to find ways of offering some grade-appropriate reasoning to make sense of (1) to fourth graders.[6] Section 7.4 of Wu, 2011, makes two such (well-known) suggestions, and we will recap one of them here. For this purpose, having a correct definition of division-with-remainder as in (5) becomes an indispensable asset. First, let us rewrite

[6] In other words, we must do some serious *mathematical engineering* to make (1) usable with a class of fourth graders. See Wu, 2006 for this concept.

(1) by putting in the zero that was intentionally omitted for simplicity:

$$\begin{array}{r@{\,}r@{\;\;}r@{\;\;}r}
 & & 1 & 9 \\ \cline{2-4}
4 &) & 7 & 8 \\
 & & 4 & \color{red}{0} \\ \cline{3-4}
 & & 3 & 8 \\
 & & 3 & 6 \\ \cline{3-4}
 & & & 2
\end{array} \tag{7}$$

This rewrite makes it obvious that the subtraction $78 - 40 = 38$ is actually an intermediate step in the long division.[7] We are now going to make some sense of (7), as follows. By the definition of the division-with-remainder of 78 by 4, we want a whole number q and a whole number r so that

$$78 = (q \times 4) + r \quad \text{where } 0 \leq r < 4 \tag{8}$$

We are going to estimate what q has to be. It cannot be a 3-digit number because the smallest 3-digit number is 100, and if q has 3-digits, then the right side of (8) ≥ 400, which would contradict (8). Next, we try letting q be a 2-digit number. If $q \geq 20$, then the right side of (8) would be ≥ 80, again impossible. So $q < 20$. Therefore, let $q = 10 + b$ where b is a single-digit number. Then $q \times 4 = 40 + 4b$. By (8), we have $78 = 40 + 4b + r$, which gives

$$\mathbf{(78 - 40)} = 4b + r \tag{9}$$

This explains the appearance of $78 - 40 = 38$ in (7). Next, we estimate what b should be. According to (9), $r = 38 - 4b$, and since $0 \leq r < 4$ by (8), we have $0 \leq 38 - 4b < 4$. At this point, a knowledge of the multiplication table immediately gives $b = 9$, so that $q = 10 + b = 19$. Thus, (9) gives $38 = 36 + r$, or

$$\mathbf{38 - 36} = r$$

On the one hand, this explains the appearance of $38 - 36 = 2$ in (7). On the other hand, we get $r = 2$. Referring back to (8), we have arrived at

$$78 = (19 \times 4) + 2$$

and this is exactly (6). We have finally made some mathematical sense of the "division house" in (7) (or (1)) as well as its kinship to (6).

[7] The omission of zeros where there is no danger of confusion is of course common practice in all the standard algorithms.

In a fourth grade or fifth grade classroom, one should use the same strategy to do a few more specific examples, e.g., why the long division of 138 by 5 leads to $138 = (27 \times 5) + 3$, or why the long division of 781 by 4 leads to $781 = (195 \times 4) + 1$. The latter example will be particularly illuminating to students because they get to see that the long division of 78 by 4 in (1) is completely imbedded in the long division of 781 by 4.

$$
\begin{array}{llccc}
 & & 1 & 9 & 5 \\ \cline{2-5}
4 &) & 7 & 8 & 1 \\
 & & 4 & & \\ \cline{3-5}
 & & 3 & 8 & \\
 & & 3 & 6 & \\ \cline{3-5}
 & & & 2 & 1 \\
 & & & 2 & 0 \\ \cline{3-5}
 & & & & 1
\end{array} \tag{10}
$$

Teaching long division in grade 6

The teaching of the long division algorithm usually spans grades 4-6. We now describe what students should learn about the algorithm by the end of the sixth grade: they should know why the long division of a two-digit number by a single-digit number—such as 78 by 4—leads *directly* to a division-with-remainder such as $78 = (19 \times 4) + 2$.

First of all, what is an *algorithm*? This word is used frequently in elementary school, yet it is hardly ever explained and even more rarely taken seriously in teaching. If the teaching of the standard algorithms would include an explicit description in each case of what the *algorithm* in question is (e.g., see Wu, 2011, pp. 63, 74, 86-87, 108-109), then the mathematical quality of the teaching would most likely improve, as we will try to demonstrate with the long division algorithm. For the purpose of school mathematics, we may define an **algorithm** to be a finite sequence of precise instructions for carrying out specific computations so that it results in a desired outcome at the end. To describe the *long division algorithm*, we should therefore write down abstractly a sequence of steps so that, for any pair of whole numbers m and n ($n \neq 0$), these steps will lead to the division-with-remainder of m by n in the form of (5). Something approximating this can be found in Section 7.3 of Wu, 2011. In a sixth grade classroom, however, such an approach would be impractical. Instead, we will describe explicitly such a finite sequence of instructions for specific cases. For example, here is the **long division algorithm of 78 by 4**. An overall comment is that each step in this sequence is a division-with-remainder whose divisor is always 4 and whose dividend will involve *one digit of the dividend* 78 *at a time.*

> **Step 1.** Perform the division-with-remainder so that its dividend is the leftmost digit 7 of 78 (recall: its divisor is always 4):
>
> $$7 = (1 \times 4) + 3 \tag{11}$$
>
> **Step 2.** Perform the division-with-remainder so that its dividend is the sum of the next digit of 78 (which is 8) and 10 times the remainder of the preceding division-with-remainder (which is 3):
>
> $$38 = (9 \times 4) + 2 \tag{12}$$
>
> (Recall: the divisor is always 4.)
>
> **Step 3.** The quotient of the division-with-remainder of 78 by 4 is obtained by "stringing together" the single-digit quotients in Steps 1 to 2, namely, 1 and 9. The remainder of the division-with-remainder of 78 by 4 is the remainder of the last step (Step 2), which is 2.

One must convince sixth graders that, strange as Steps 1-3 may seem, the long division in (1), upon closer inspection, is nothing more than a schematic representation of Steps 1 and 2. What we want to show is that *the long division algorithm is correct*, i.e., we have to prove the following theorem.

Theorem 1. *Steps 1 and 2 imply Step 3.*

One may think that Theorem 1 is a waste of time, because to show Step 3 is correct, all we have to do is check that $78 = (19 \times 4) + 2$ is correct. But the theorem says much more: it says that Step 3 *can be derived strictly from Steps 1 and 2.* Thus it is more than a numerical statement that $(19 \times 4) + 2$ is equal to 78. Rather, it asserts that we can use reasoning alone to get to the equality $78 = (19 \times 4) + 2$ by making use of Steps 1 and 2.

Proof of Theorem 1. As in the proofs of the validity of all the standard algorithms, the key ingredient is the *expanded form of a whole number* (see, e.g., page 20 of Wu, 2011):

$$78 = 70 + 8 \tag{13}$$

Now, from (11), we get $7 = (1 \times 4) + 3$. Therefore,

$$70 \;=\; (10 \times 4) + 30$$

Substituting this value of 70 into (13), we get $78 = (10 \times 4) + 30 + 8$, which is equal to

$$78 \;=\; (10 \times 4) + 38$$

(Observe that this corresponds to the subtraction $78 - 40 = 38$ in the "division house" (1).) Substituting the value of 38 in (12) into the right side of the preceding equation, we obtain

$$78 \;=\; (10 \times 4) + (9 \times 4) + 2$$

Applying the distributive law to the first two terms on the right side, we get

$$78 \;=\; (19 \times 4) + 2$$

This shows that the division-with-remainder of 78 by 4 has quotient 19 and remainder 2, exactly as claimed by Step 3. Theorem 1 has been proved.

The first question we must ask is in what way are this theorem and its proof superior to the above informal argument presented in connection with equations (8) and (9). The answer is that insofar as the long division algorithm is an *algorithm*, we are duty bound to give an explicit description of every step of the algorithm, and this the earlier informal argument failed to do. Precision and clarity matter in mathematics. Moreover, mathematics is about the deduction of conclusions from assumptions, and the preceding theorem and its proof present a textbook case of this deduction process. By comparison, one is left uncertain about the precise assumptions that were made in the earlier argument. Also see the comment following the third remark below.

The preceding proof should also be supplemented by three additional remarks. First, the long division algorithm exemplifies the recurrent theme of the standard algorithms, which is to *break up a multi-digit computation into computations involving single digits* (see Chapter 3 of Wu, 2011). Thus each of Steps 1 and 2 essentially (though not literally) computes with the digits of the dividend 78 *one at a time*, and more importantly, the algorithm itself computes the quotient 19 *one digit at a time* (see Step 3). It may also be observed that although each of Steps 1 and 2 is itself a division-with-remainder, it differs from the original division-with-remainder of 78 by 4 in that the dividend in each of Step

1 and Step 2 (7 and 38, respectively) is smaller than the original dividend of 78. While this fact may not seem to be much of an advantage when the original dividend (such as 78) is relatively small, the advantage will become more pronounced as the dividend gets larger. Our next example of the long division of 781 by 4 will give a better idea in this regard.

A second remark is that, to the extent that there should be one general *long division algorithm* that is applicable in all cases, one may not be able to discern from the preceding Steps 1-3 what the general long division algorithm should look like. This lack of clarity will disappear, however, in our next two examples with a dividend of three digits. For a more precise description of the general case, see Chapter 7 of Wu, 2011.

A third remark is that one should take note of the fact that the algorithm, as stated in Steps 1-3, completely ignores the place value of the digits of the dividend. This fact will be more forcefully brought out after we discuss the division-with-remainder of 781 by 4. Contrary to the emphasis placed by the education literature on the concept of place value in discussing the standard algorithms, a main selling point of the standard algorithms is the *mathematical simplicity* of their execution because these algorithms intentionally ignore place value (see Wu, 2011, pp. 59, 66, 120-121). Place value becomes relevant only when we try to *prove* that an algorithm is correct. From this perspective, the argument in connection with equations (8) and (9) is unsatisfactory because it does not draw a sharp line between the place-value independence of the algorithm itself and the key role place value plays in the justification of the algorithm.

We should also mention that there is a subtle issue involving the implicit assumption in Steps 1 and 2 that the quotient in each division-with-remainder of (11) and (12) will be a *single-digit* number. We refer the reader to Section 7.6 of Wu, 2011 for the simple explanation.

As promised, we will next take up the division-with-remainder of 781 by 4. In a typical sixth grade classroom, this example would be optional, though highly desirable.

Recall first of all that no thinking is needed for getting the division-with-remainder of 781 by 4: count all the multiples of 4 up to 781. However, this is clearly a tedious process and a shortcut is called for.[8] As we mentioned earlier, the long division algorithm is the

[8]The tedium would be more obvious if we contemplate the division-with-remainder of 789123 by 43.

sought-for shortcut. Let us first recall the long division in (10):

$$\begin{array}{r|lll} & 1 & 9 & 5 \\ \hline 4 \;) & 7 & 8 & 1 \\ & 4 & & \\ \cline{2-4} & 3 & 8 & \\ & 3 & 6 & \\ \cline{2-4} & & 2 & 1 \\ & & 2 & 0 \\ \cline{2-4} & & & 1 \end{array} \tag{14}$$

It is easy to verify that this "division house" is merely a schematic representation of Steps 1-3 of the following **long division algorithm of 781 by 4**:

> **Step 1.** Perform the division-with-remainder so that its dividend is the leftmost digit 7 of 781 (recall: its divisor is always 4):
>
> $$7 = (1 \times 4) + 3 \tag{15}$$
>
> **Step 2.** Perform the division-with-remainder so that its dividend is the sum of the next digit of 781 (which is 8) and 10 times the remainder of the preceding division-with-remainder (which is 3):
>
> $$38 = (9 \times 4) + 2 \tag{16}$$
>
> (Recall: the divisor is always 4.)
>
> **Step 3.** Perform the division-with-remainder so that its dividend is the sum of the next digit of 781 (which is 1) and 10 times the remainder of the preceding division-with-remainder (which is 2):
>
> $$21 = (5 \times 4) + 1 \tag{17}$$
>
> (Recall: the divisor is always 4.)
>
> **Step 4.** The quotient of the division-with-remainder of 781 by 4 is obtained by "stringing together" the single-digit quotients in Steps 1-3, namely, 1, 9, and 5. The remainder of the division-with-remainder of 781 by 4 is the remainder of the last step (Step 3), which is 1.

What we want to prove is that the long division algorithm of 781 by 4 is correct, i.e., we have the following theorem.

Theorem 2. *The preceding Steps 1-3 imply Step 4.*

Proof. Having gone through the proof of Theorem 1 in detail, we will be more brief this time around. As always, we begin with the expanded form of 781:

$$781 = 700 + 80 + 1 \tag{18}$$

From (15), we get

$$700 = (100 \times 4) + 300$$

Substituting this value of 700 into (18), we have $781 = (100 \times 4) + 300 + 80 + 1$, or

$$781 = (100 \times 4) + 380 + 1 \tag{19}$$

Now (16) implies that $380 = (90 \times 4) + 20$. If we substitute this value of 380 into (19), we obtain

$$781 = (100 \times 4) + (90 \times 4) + 20 + 1$$

Applying the distributive law to the first two terms on the right side, we get

$$781 = (190 \times 4) + 21$$

Now substituting the value of 21 in (17) into the right side, we obtain

$$781 = (190 \times 4) + (5 \times 4) + 1$$

Using the distributive law again on the right side, we finally arrive at

$$781 = (195 \times 4) + 1 \tag{20}$$

Since this is exactly the statement of Step 4 above, i.e., the division-with-remainder of 781 by 4 has quotient 195 and remainder 1, the proof of the theorem is complete.

Remarks. (1) Now it should be clear from the repetitive nature of the preceding Steps 2 and 3 how the long division algorithm will proceed in the general case: Begin with the leftmost digit of the dividend as in Step 1 above and repeat the following process until you get to the rightmost digit of the dividend:

> For the dividend of the next division-with-remainder, add 10 times the remainder of the preceding division-with-remainder to the next digit to the right in the original dividend.

Moreover, it is equally clear how to prove that the algorithm is correct: start with the expanded form of the original dividend and replace each term in the expanded form by each of the divisions-with-remainder given by the steps of the algorithm.

(2) Looking back over our work so far, we can see more clearly the purpose of the long division algorithm: it is to replace the original division-with-remainder by a succession of simpler divisions-with-remainder in each of which the dividend is smaller than the original one. Thus, in the case of the division-with-remainder of 781 by 4, the dividends in (15)-(17) are 7, 38, and 21; each is far smaller than 781.

(3) We are also in a better position now to understand the statement that the long division algorithm *ignores place value*. Let us compare the two long divisions: 78 by 4 and 781 by 4. The number 7 is the tens digit in 78 but is the hundreds digit in 781, yet the first steps of the algorithm in the two cases, (11) and (15), are identical. Similarly, the number 8 is the ones digit in 78 but is the tens digit in 781, and yet the second steps of the algorithm in the two cases, (12) and (16), are again identical. These confirm a key fact about the long division algorithm: it only looks at each individual digit of the dividend but not its place value. (Let it be said one more time that the proof of the validity of the algorithm *does* take into account the place value of each digit of the dividend.)

To consolidate our gains, we will take up the division-with-remainder of 242 by 16 (this is the division-with-remainder suggested in Green, 2014). Again, in a typical sixth grade classroom, this would be optional, though extremely instructive. The new feature here is that the divisor is a two-digit number. In this case, the **long division algorithm of 242 by 16** is the following:

> **Step 1.** Perform the division-with-remainder so that its dividend is the leftmost digit 2 of 242 (recall: its divisor is always 16):
>
> $$2 = (0 \times 16) + 2 \tag{21}$$
>
> **Step 2.** Perform the division-with-remainder so that its dividend is the sum of the next digit of 242 (which is 4) and 10 times the remainder of the preceding division-with-remainder (which is 2):
>
> $$24 = (1 \times 16) + 8 \tag{22}$$

(Recall: the divisor is always 16.)

Step 3. Perform the division-with-remainder so that its dividend is the sum of the next digit of 242 (which is 2) and 10 times the remainder of the preceding division-with-remainder (which is 8):

$$82 \;=\; (5 \times 16) + 2 \tag{23}$$

(Recall: the divisor is always 16.)

Step 4. The quotient of the division-with-remainder of 242 by 16 is obtained by "stringing together" the single-digit quotients in Steps 1-3, namely, 0, 1, and 5. The remainder of the division-with-remainder of 242 by 16 is the remainder of the last step (Step 3), which is 2.

Here is the "division house" of the long division of 242 by 16:

$$\begin{array}{ccccccc}
 & & & 0 & 1 & 5 \\ \cline{4-6}
1 & 6 &) & 2 & 4 & 2 \\
 & & & 0 & & \\ \cline{4-5}
 & & & 2 & 4 & \\
 & & & 1 & 6 & \\ \cline{4-6}
 & & & & 8 & 2 \\
 & & & & 8 & 0 \\ \cline{4-6}
 & & & & & 2
\end{array} \tag{24}$$

It is easy to see that this "division house" is nothing but a schematic presentation of the preceding Steps 1-3. Let us prove once again that the algorithm is correct.

Theorem 3. *In the long division algorithm of 242 by 16, Steps 1-3 imply Step 4.*

Proof. The expanded form of 242 reads:

$$242 \;=\; 200 + 40 + 2 \;=\; 240 + 2 \tag{25}$$

Since (21) is the trivial statement that $2 = 2$, we begin with (22), which implies that $240 = (10 \times 16) + 80$. Substituting this value of 240 into (25), we get

$$242 \;=\; (10 \times 16) + 80 + 2 \;=\; (10 \times 16) + 82$$

Now substituting the value of 82 in (23) into the preceding equation, we obtain

$$242 \;=\; (10 \times 16) + (5 \times 16) + 2$$

Applying the distributive law to the first two terms on the right side, we get

$$242 \;=\; (15 \times 16) + 2$$

which is exactly the statement that the division-with-remainder of 242 by 16 has quotient 15 and remainder 2, i.e., Step 4 is correct. The theorem is proved.

Again, the repetitive nature of Steps 2 and 3 helps to give a clear conception of what the general long division algorithm is about. This *long division algorithm of 242 by 16* also serves to better highlight a special feature of the long division algorithm in general, which is to break up the original division-with-remainder of 242 by 16 into more manageable divisions-with-remainder, each with the same divisor 16 but with a far smaller dividend: the division-with-remainder of 24 by 16, and the division-with-remainder of 82 by 16. One more thing that is noteworthy is that TSM teaches the long division of 242 by 16 by saying that, since 16 does not go into 2, one should consider the first two digits 24 of 242 as the first dividend. However, the long division algorithm of 242 by 16—being an *algorithm*—does not depend on this *contingent* kind of judgment to "skip a step" in certain situations. Its instruction to perform a division-with-remainder in Step 1 is meant to be carried out literally, as it was in (21).

Curricular implications

Because of the lack of space, we will be brief in giving an indication how TSM's mangling of the concept of division-with-remainder has pernicious repercussions later in the school mathematics curriculum.

The concept of the **gcd** (greatest common divisor)[9] of two nonzero whole numbers is a staple of elementary school mathematics, but TSM's failure to correctly teach division-with-remainder has forced the teaching of gcd to be confined entirely to an inspection of the factors of each number. In particular, this failure results in the Euclidean algorithm not being taught in K-12 as an effective method of getting the gcd. While we are not strongly advocating here that the Euclidean algorithm be taught in K-12, we can nevertheless amplify on the fact that, by not teaching division-with-remainder properly, TSM hampers students' future mathematics learning. Consider, for example, a favorite

[9]Highest Common Factor (HCF) in school mathematics.

activity in the learning of fractions: how to simplify the following fraction to lowest terms:

$$\frac{551}{247}$$

It is not so easy to factor either 551 or 247, but if we do the long division of 551 by 247, we get

$$551 \;=\; (2 \times 247) + 57$$

It is a fairly straightforward consequence of this equality that the two pairs $\{551, 247\}$ and $\{247, 57\}$ have exactly the same collection of common divisors (see page 465 in Wu, 2011 or page 210 in Wu, 2016a). But the second pair has the advantage that 57 is considerably smaller than 247 or 551 and therefore the search for the gcd of $\{247, 57\}$ promises to be potentially easier than that of $\{551, 247\}$. As a matter of fact, it is so easy to get all the factors of 57 (they are 1, 3, 19, and 57 because $57 = 3 \times 19$) that we know the common divisors of $\{247, 57\}$ are among 1, 3, 19, and 57. It is then painless to conclude that the gcd of $\{247, 57\}$ is in fact 19 so that 19 is also the gcd of 551 and 247. Hence,

$$\frac{551}{247} \;=\; \frac{19 \times 29}{19 \times 13} \;=\; \frac{29}{13}$$

The idea that the task of finding the gcd of two whole numbers can be simplified by just one application of long division—without a doubt—deserves to be taught, but it *cannot be taught* if all that TSM has to offer about long division is of the "$551 \div 247 = 2\,R57$" variety.

Needless to say, the Euclidean algorithm is just an iteration of the preceding process (see, e.g., pp. 464ff. in Wu, 2011 or pp. 203ff. in Wu, 2016a).

Of course, the numbers in the preceding problem have been rigged to heighten the drama(!), but the message should not be lost, to the effect that the long division algorithm is not just a boring arithmetic skill but is a very versatile mathematical tool (see, e.g., Part 4 of Wu, 2011). More importantly, such an application of the long division algorithm is a powerful illustration of the coherence of mathematics: the fact that there are hidden connections between seemingly disparate topics, e.g., long division and simplifying fractions. TSM routinely makes it impossible for students to see this and other hidden connections.

The issue of the coherence of mathematics naturally brings us to two other topics in the school curriculum related to long division. First, the **division algorithm** for polynomials of one variable states that given two such polynomials $F(x)$ and $G(x)$ (with

$G \neq 0$), F can be expressed in terms of G and two polynomials $Q(x)$ and $R(x)$ as follows:

$$F(x) \;=\; Q(x)G(x) + R(x), \quad \text{where } \deg R(x) < \deg G(x)$$

(see, e.g., Section 5.1 of Wu, 2020b). This is clearly an analogy of division-with-remainder with Q and R playing the roles of **quotient** and **remainder**, respectively, and F and G playing the roles of **dividend** and **divisor**, respectively. This is not quite a direct generalization from whole numbers to polynomials because, whereas for the division-with-remainder, the comparison of the remainder to the divisor is by using the magnitudes of the whole numbers, the comparison in the case of the division algorithm for polynomials is by using the *degrees* of the polynomials. Thus polynomial division is reminiscent of long division but not a generalization of it. In advanced mathematics, both will become special cases of *Euclidean domains*—again, a reminder of the coherence of mathematics. But since TSM does not teach division-with-remainder, there can be no such intellectual resonance when students come to the division algorithm for polynomials in Algebra II. This is an opportunity wasted.

Our next topic is about the conversion of a fraction to a decimal "by the long division of the numerator by the denominator". This is such a well-known topic in middle school mathematics that it suffices to use a simple example to illustrate the fundamental ideas involved. We claim:

$$\frac{7}{16} \;=\; 0.4375 \tag{26}$$

where 4375 is the quotient of the *long division* of 7×10^4 by 16. Here, the validity of the equality of the two numbers in (26) is not in question because, by the definition of a decimal (see Wu, 2011, page 187),

$$0.4375 \;=\; \frac{4375}{10,000}$$

and the fraction on the right easily simplifies to $\frac{7}{16}$. What is at issue, rather, is why the fraction $\frac{7}{16}$ is *equal to* the quotient of the long division of 7×10^4 by 16 with the decimal point inserted in front. (Just as in Theorem 1, we have again an example of the emphasis placed on the *method* of arriving at a conclusion rather than on the validity of the conclusion itself.) This issue—that the fraction is *equal* to the decimal produced by the long division—is precisely what is not addressed in the educational discussions in TSM of this basic conversion procedure.[10] This is not surprising because we have seen

[10]Such discussions are usually preoccupied with the "repeating" property of the decimal.

the pattern in TSM of interpreting the equal sign in (26) as a command to *compute by the long division of 7 by 16 to get the decimal 0.4375.* Therefore, what we seek is not just a proof of (26) *per se*, but an explanation for the intrusion of long division into this conversion procedure and how the decimal point materializes. To this end, we will prove something more general: let the *long division* of 7×10^k (k being any nonzero whole number) by 16 have quotient N and remainder r, then

$$\frac{7}{16} = \frac{N}{10^k} + \frac{r}{16 \times 10^k} \tag{27}$$

First, we show that (27) implies (26). Let $k = 4$ in (27). Since $7 \times 10^4 = 4375 \times 16$, we see that $N = 4375$ and $r = 0$ in this case. Therefore (26) immediately follows. Next, to prove (27), we have by the definition of N and r that

$$7 \times 10^k = (N \times 16) + r$$

Therefore,

$$\frac{7}{16} = \frac{7 \times 10^k}{16 \times 10^k} = \frac{(N \times 16) + r}{16} \times \frac{1}{10^k} = \frac{N}{10^k} + \frac{r}{16 \times 10^k}$$

The proof of (27) is complete. We also pause to note that, included in (27) is the assertion that if we let $k = 5$ or any whole number > 4, (27) will still imply (26).

It is not out of place to remark that the simplicity of the preceding reasoning comes from the clear understanding of how long division leads to division-with-remainder as in Theorems 1-3. We should also point out that (27)—together with its proof above—remains valid, verbatim, when the fraction $\frac{7}{16}$ is replaced by *any* fraction. This is the key step that leads to the general proof of the fact that the conversion of a fraction to a decimal can be achieved "by the long division of the numerator by the denominator". For the details, see Section 3.4 of Wu, 2020c. (A simplified version without the use of limits has been given in Chapter 42 of Wu, 2011.)

The conversion of a fraction to a decimal by long division is a staple of school mathematics, yet it seems almost impossible to locate a correct proof in the recent education literature.[11] So long as TSM rules over school mathematics and so long as it preaches that $16 \div 7 = 2\,R2$, there can be no hope for such a proof. Once again, the surprising connection between long division and the conversion of a fraction to a decimal is lost because TSM has rendered mathematical reasoning impossible.

[11] It is taken for granted that someone must have written a correct proof sometime in the distant past.

Must we eradicate TSM?

In her popular article (Green, 2014), Elizabeth Green makes a passing comment on the traditional way of teaching the "division house" as nothing more than a ritualistic inculcation of mind-numbing procedures that turns "school math into a sort of arbitrary process wholly divorced from the real world of numbers". Green's article suggests that the road to improvement is to change the *teaching* of school mathematics by tapping "into what students already understood and then [building] on it.. . . By pushing students to talk about math", they will uncover their own misunderstandings about division and make sense of the "division house".

We prefer to take a simpler, more grounded view on the rampant non-learning of school mathematics in the past decades. When teachers are taught only TSM, they will naturally teach only TSM to their students. The ritualistic inculcation of procedures in school mathematics classrooms is mostly—though not totally—a reflection of teachers' content-knowledge deficit. But teachers are not to blame for the sorry spectacle—the education establishment is. Why should we expect teachers to help students understand school mathematics when we have made no effort to help *teachers* understand it? It is altogether unrealistic to expect teachers to uncover by themselves the substantial mathematical content that undergirds the "division house" (see Theorem 1-3 above). All of us in the education community share the responsibility for helping teachers shed their baggage of TSM and acquire a new and *correct* knowledge base for teaching, but on this job, we have thus far fallen flat on our faces. Once we succeed in providing teachers with the *mathematical knowledge they need for teaching*, then it would be realistic to consider how teachers could better *teach* school mathematics.

The main contention of this article is that TSM is destroying school mathematics education. Now, one does not make such a sweeping statement unless one has incontrovertible proof. In this case, the proof of how destructive TSM really is can be easily accessed everywhere: pick up any of the standard school mathematics textbooks published between 1970 and 2015 or, in all likelihood, if you pick up any of the current textbooks and open it up to a random page, your chance of witnessing TSM at work is likely to exceed 80%. It is not possible in year 2020 to have a meaningful discourse on improving school mathematics education without first initiating some serious attempts to get rid of TSM.

This article has chosen to focus on one topic, long division, to reveal the truly antimathematical nature of TSM. In our extended discussion, we have shown how TSM has perverted the mathematics of this single topic and made it unlearnable. But of course, we can say the same about almost every topic in the mathematics of K-12. If we want to change course and make school mathematics learnable, we must have the necessary political will to make the change and the political muscle to implement this change. But what has been long overlooked is that, besides the will and the muscle, there must also be a detailed version of *correct and learnable* school mathematics from K to 12—*all of it*—that is universally available to provide guidance. It is the absence of such an alternate version that has shipwrecked every reform since the 1950's (see pp. 7-10 of Wu, 2020d). The author has just completed a first attempt at presenting such an alternate version in the form of six volumes: Wu, 2011 (for elementary school), Wu 2016a and Wu 2016b (for middle school), and Wu 2020a, Wu 2020b, and Wu 2020c (for high school). In such a large undertaking (the six volumes comprise about 2,500 pages), some topics, no matter how significant, will inevitably get short shrift, and the long division algorithm is one such example. This article may therefore be regarded as one of many needed supplementary commentaries on these six volumes. We hope that it will also help to raise awareness of how pervasive has been the damage done by TSM.

It is quite common to hear people say "I am not good at math" without a trace of self-consciousness or embarrassment. What they actually *mean* is something like "I am no good at memorizing an unending collection of meaningless factoids and bags of tricks for getting answers (i.e., TSM)". Indeed, no one should feel self-conscious or embarrassed about not being able to learn TSM. The urgent question is: can we spare the next generation the trauma—and the inevitable punishment for failure—of a thirteen-year immersion in unlearnable math?

References

[1] Falkner, K. P., Levi, L., and Carpenter, T. P. (1999). Children's understanding of equality: A foundation for algebra. *Teaching Children Mathematics*, 6, 232-236. `https://tinyurl.com/y4a7q3vv`

[2] Green, E. (2014) Why do Americans stink at math? The New York Times Magazine, July 27. `https://tinyurl.com/s724drd`

[3] Kieran, C. (1981). Concepts associated with the equality symbol. *Education Studies in Mathematics*. Vol. 12, 317-326.

[4] Knuth, E., Alibali, M. W., Hattikudur, S., McNeil, N. M., and Stephens, A. C. (2008). The importance of equal sign understanding in the middle grades, *Mathematics Teaching in the Middle School*. Vol. 13, 514-519.

[5] Wu, H. (2006). How mathematicians can contribute to K-12 mathematics education, *Proceedings of International Congress of Mathematicians, Madrid 2006*, Volume III. Zürich: European Mathematical Society, 1676-1688. `http://math.berkeley.edu/~wu/ICMtalk.pdf`

[6] Wu, H. (2011). *Understanding Numbers in Elementary School Mathematics.* Providence, RI: American Mathematical Society.

[7] Wu, H. (2016a). *Teaching School Mathematics: Pre-Algebra.* Providence, RI: American Mathematical Society.
Its ***Index*** *is available at:* `http://tinyurl.com/zjugvl4`

[8] Wu, H. (2016b). *Teaching School Mathematics: Algebra.* Providence, RI: American Mathematical Society.
Its ***Index*** *is available at:* `http://tinyurl.com/haho2v6`

[9] Wu, H. (2020a). *Rational Numbers to Linear Equations*, Providence, RI: American Mathematical Society.

[10] Wu, H. (2020b). *Algebra and Geometry*, Providence, RI: American Mathematical Society.

[11] Wu, H. (2020c). *Pre-Calculus, Calculus, and Beyond*, Providence, RI: American Mathematical Society.

[12] Wu, H. (2020d). What can we do about our dysfunctional school mathematics curriculum? `https://math.berkeley.edu/~wu/RNLE1.pdf`

Department of Mathematics
University of California at Berkeley