A SCIENCE-OF-LEARNING APPROACH TO MATHEMATICS EDUCATION

FRANK QUINN

Abstract. A scientifically disciplined, learning-oriented approach to mathematics education is illustrated through examples and detailed comparison to articles in the March 2011 Special Issue on Education of the *Notices* of the American Mathematical Society. Many of the *Notices* articles focus on teacher education, with an underlying presumption that the current teacher corps is not sufficiently competent. The analysis here suggests that the teachers are competent enough, but the methodology they have been taught to use may not be competent. In other words, significantly better outcomes may require profound changes in educational philosophy and teacher-education programs.

Contents

Date: April 13 2011.

1. Introduction

This article[1] gives a perspective on K-12 mathematics education that differs from standard approaches in several ways. First, the focus is directly on individual learning rather than on teaching, and conclusions about teaching are driven by learning considerations. Second, the perspective developed bottom-up from micro-scale observations. It owes much to cognitive neuroscience, but little to educational philosophy since this has little traction at micro scales. Third, it developed top-down in terms of educational level. Our eventual goal is to produce competent citizens and capable scientists and engineers. To do this, college faculty need certain skills from high school; to develop these skills, high-school teachers need certain outcomes from middle school; and so on. In particular, long-term needs impose strong constraints on elementary education. Finally the account draws on modern professional practice. These features are described a bit more in the next section, and in the preface to [7].

Much of the output from this approach is also micro-scale: highly-effective ways to treat fractions, polynomial multiplication, word problems, etc., but not much about how these might be put together in a course or curriculum. In other words, a toolkit without assembly instructions. Section 3 illustrates this through examples. Topics are fractions, word problems, multiplication of polynomials, and partial fractions in both integers and polynomials. The examples are middle- and high-school topics because the issues are clearer and the alternatives richer, but many of the conclusions apply to lower grades as well. Many of the examples are taken from [1], and comparisons are made in the final section.

Section 4 describes potential implications for teaching. There are several points related to the description as “a toolkit without assembly instructions”. First, the teaching implications are mainly strategies to make the tools work, so do not address “assembly” issues such organizing or presenting the material. This is in contrast to educational-philosophy approaches that provide assembly instructions without much toolkit, and focus more on pedagogy and content-independent methodology. The second point is that the suggestions are relevant only for people who want to use the tools effectively. They are not method-independent assertions about teaching.

The second subsection directly concerns teaching. The first provides a coherent context: a description of mathematics developed to help make sense of mathematical aspects of common features. The final section concerns teacher preparation relevant to this approach. Among other things it throws light on why study of advanced topics does not seem to improve K-12 teaching.

The final Section, 5, contrasts some of these conclusions with articles in March 2011 special issue on education of the *Notices* of the American Mathematical Society, [1]

2. Background

The most significant background for this material is extensive (over a thousand hours) one-on-one diagnostic work with students[2]. The procedure is that I go to them when they need help. I review their work, diagnose the specific problem, and

[1]This is the full version. The abridged version omits some examples and the final section on comparisons with [1].

[2]In the Math Emporium at Virginia Tech.

show them how to fix it to avoid a recurrence.This is not tutoring. The students do most of the talking and most of the work, and I leave to let them finish on their own as soon as they are past the specific difficulty. The focus on diagnosis also helps students see how to diagnose and correct their own errors.

The context for the diagnostic work is a computer-tested engineering calculus course that I developed. Diagnosis of problems provides feedback for the courseware: it is designed to emphasize structure and encourage work habits that avoid these problems. The material is also designed so that pre-existing learning errors will cause serious difficulty. The goal is to expose these errors so they can be diagnosed and repaired.

A final ingredient is an extensive bottom-up analysis of contemporary mathematical practice, [6]. This draws on my own research experience, editorial work, work on publication policy, cognitive neuroscience, history of mathematics, the work with students mentioned above, and many other sources. It is well-known that modern professional practice is quite different from that in the nineteenth century, and from models used in education and most philosophical accounts. It was no surprise to find that contemporary practice is better adapted to the subject. It was unexpected to find that, within the subject constraints, it is much better adapted to human cognitive facilities. It seems that contemporary practice is not just for freaks: carefully understood, it can be a powerful resource for education.

These experiences have given a fine-grained perspective on mathematical learning and its difficulties. Moreover the perspective directly concerns individual learning: not teaching, and not learning mediated by teachers. I have almost forty years of traditional classroom teaching experience, and found—to my dismay—that teaching is not nearly as closely connected with learning as I had believed.

3. Examples

Explicit problem-oriented methods for student use are the main products of the science-of-learning approach and correspondingly are the core of this essay. But isolated novelties are of limited value, so enough of these must be given so they can be seen to fit together. An unconventional approach to fractions in the first section, for example, is shown to give effective contexts for partial fractions in both the integers (§3.3) and polynomials (§3.3). Similarly an unconventional procedure for multiplying polynomials in §3.4 is adapted to find coefficients in polynomial partial fractions in §3.3, and a modular-arithmetic approach to a tricky word problem in §3.2 is related to partial fractions in both integers and polynomials. Many more examples and interconnections are given in [7]. The polynomial multiplication procedure, for instance, is adapted to give an algorithm for multiplying multi-digit integers in *Proof projects for teachers* and *Neuroscience experiments for mathematics education.*

3.1. **Fractions.** These are a perennial source of trouble, and there seem to be two reasons. First, a fraction is a name for something, not the thing itself. Of course all our symbols are names, not things. This point is a bit subtle for elementary education, and if only one name is used then imprecision ("2 is a number") seems harmless. But the expression $\frac{9}{4} = 2.25$ really means "these are both names for the same thing". The different names encode different properties, and we want to work with both because we want to exploit the different information they encode.

Identifying the names with the number makes this obscure and confusing. There is further discussion of names in §4.2.2.

The second point about fractions is that they specify things implicitly. The name 2.25 encodes an explicit procedure for assembling a number from single-digit integers and powers of 10. Fractions and names such as $\sqrt{2}$ encode properties that determine the thing, but do not encode a procedure. See the comments below, and in §4.2.2.

3.1.1. *Definition.* A *faction* is a name that encodes a property. Specifically, $\square = \frac{a}{b}$ means that $\square$ has the property $b\square = a$.

3.1.2. *Examples.*

(1) The decimal 2.25 is entitled to the name $\frac{9}{4}$ because it satisfies $4 \times 2.25 = 9$.
(2) Is $a - 2 \stackrel{?}{=} \frac{a^2-4}{a+2}$? To find out, see if it has the property $(a+2) \times \square \stackrel{?}{=} a^2 - 4$. Trying it gives
$$(a+2) \times (a-2) = a^2 - a2 + 2a - 4 = a^2 - 4,$$
so the answer is 'yes'.
(3) Is $\frac{a}{x} + \frac{b}{y} \stackrel{?}{=} \frac{a+b}{x+y}$? To see if the defining property is true, clear denominators (multiply by $xy(x+y)$). We get $a(y(x+y)) + b(x(x+y)) \stackrel{?}{=} (a+b)(xy)$. Expanding and canceling gives $ay^2 + bx^2 \stackrel{?}{=} 0$. This hardly ever works, so the answer in general is 'no'.
(4) Express $\frac{a}{x} + \frac{b}{y}$ as a fraction. Plan: first give this a neutral name, $\square = \frac{a}{x} + \frac{b}{y}$. Expressing it as a fraction, $\square = \frac{r}{s}$, means we want to find r, s so that $s\square = r$. The original expression is given in terms of fractions so to learn anything about it we have to clear denominators and use the defining property. Multiply by xy to get $xy\square = xy(\frac{a}{x} + \frac{b}{y}) = ya + xb$. But comparing with what we want shows that $\square = \frac{ya+xb}{xy}$

3.1.3. *Notation.* I explain the reasons for nonstandard notations $\square$ and $\stackrel{?}{=}$.

First, $\square$. If we are trying to find a fraction name for something, we need another name to use during the process. The thing is usually given with a compound name (e.g. $\frac{2}{a} + 7 \times \frac{2}{b}$) but it is hard to work with this and still think of it as a single object. It helps to use a temporary name, e.g "let $A = \frac{2}{a} + 7 \times \frac{2}{b}$". But using a symbol means we have *three* symbolic names in play. Using a neutral placeholder, $\square$, seems to be less confusing. This is obviously temporary, encodes no structure, and has no hidden significance.

The notation $\stackrel{?}{=}$ addresses problems with very sloppy use of '='. In common use this can mean "define x by…"; "it *is always* true that…"; "*is it* always true that…?", or "determine *when* it is true that…". The equation alone is incomplete and the intended meaning is to be inferred from context . This ambiguity becomes invisible to experts, but students—and many teachers—do not understand it and it is a great source of confusion (see question 2, p. 390, in [5]). Computers don't understand it either, and the programming language of *Mathematica* provides at least six different symbols for different meanings of '='. In the example here we want to begin with an equality posed as a question, and end with the same expression as an assertion. Confusion is inevitable if the notation does not display the difference between the two.

3.1.4. *Comments.* The 'property-encoding name' point of view dissipates much of the confusion about fractions. Sure the rules are different. If we are combining things with a special properties (encoded in the names), and want to see that these properties give the outcome a similar structure, then we use manipulations appropriate to the property. For example we could add integer fractions by converting them to decimals and adding these. This would give the right *number*, but would not display it as a fraction. The same is true for square roots, or for that matter, decimals. $1.23 + 45.6$ already specifies a number, and decimal arithmetic is necessary to find the decimal *name* for this number, not to find the number.

The precise definition also clarifies the status of the strange rules. According to the definition, if you want to do anything with fractions you have to begin by clearing denominators. The rules provide shortcuts. For instance "divide by inverting and multiplying" is not pulled out of some dark place at random, but conveniently summarizes the outcome of a longer calculation that uses the definition.

3.2. **Word problems.** This is a crucial topic because contemporary educators put great emphasis on word problems. The contemporary approach is problematic, however. There is an abstract discussion in §4.2.4; here I illustrate the use of modeling to avoid the problem.

A *mathematical model* is a translation of a real-world or word problem to a symbolic form suitable for mathematical analysis. The key feature is that little or no analysis is done during modeling, and no modeling is done during analysis of the model. Five examples are given. The first comes from the article of Heaton and Lewis [4] in the *Notices* special issue, and the approach used here is contrasted with theirs in §5.3. The second is a typical number-and-word problem with little mathematical content. The third illustrates the use of algebra to explore a problem about statistics. The fourth and fifth illustrate deficiencies of "trick problems".

3.2.1. *Example 1: Chicken nuggets.* The "chicken nugget conundrum[3]" is:

> Chicken nuggets are available in three size boxes: six, nine, and twenty. What is the smallest number of nuggets that you *cannot* get by ordering combinations of these three sizes?

3.2.2. *Mathematical model.* The mathematical model for the chicken nugget problem is:

> What is the smallest integer that *cannot* be expressed as $6a + 9b + 20c$, with a, b, c nonnegative integers?

Note that this is a direct translation, without any mathematical processing at all. The work from now on is entirely mathematical, with no cognitive overhead or confusion about food or boxes.

The model makes the mathematical issues much clearer. The greatest common divisor of 6, 9, and 20 is 1 so *any* integer can be expressed as combination if negative coefficients are allowed. The restriction on realization comes from the restriction to *nonnegative* coefficients. The assumption is not explicit in the problem so it is easy to miss, especially if the problem is not modeled. Someone impressed with

[3] A 'puzzler' from the National Public Radio program *Car Talk*.

the "debit" explanation of negative numbers might even question its validity: "can one get 3 nuggets by getting a box of 9 and an empty 6-nugget box, putting 6 in the empty box and selling it back to the restaurant? You did not *say* this is not allowed." In any case this makes the question more subtle than a congruence problem. Another complication is that primes 2, 3, and 5 are involved, and both of the smaller ones appear in two of the given numbers.

3.2.3. *Solution.* A solution is presented for future reference. Set $k = 6a + 9b + 20c$. We want to estimate the coefficients in terms of information about k.

Modular arithmetic should be the obvious approach. To get information about the coefficient c, work modulo the greatest common divisor of the other two denominators, namely 3. Mod 3 the equation becomes $k \overset{3}{\equiv} 20c \overset{3}{\equiv} 2c$. Multiplying by 2 (the inverse of 2, mod 3) gives $2k \overset{3}{\equiv} c$. This determines c mod 3, and since 3×20 can be expressed as a (nonnegative) combination of 6 and 9, we can assume c is 0, 1, or 2. This gives the first conclusion: if $k \overset{3}{\equiv} 20c$ for $0 \leq c < 3$, then $k \geq 20c$.

Next suppose $k \overset{3}{\equiv} 20c$ with $c = 0$, 1, or 2, and $k \geq 20c$. Then $k - 20c = 3m \geq 0$, and k can be realized if and only if $3m$ can be realized as a nonnegative combination of 6 and 9. Or equivalently, if m can be realized as $2a + 3b$ for nonnegative a, b. This is a smaller version of the original problem, and can be done the same way (i.e. reduce mod 2 or 3 to relate m and one of the coefficients). Eventually we see that 1 is the only nonnegative integer that cannot be realized in the smaller problem, so the largest in the original problem is $2 \times 20 + 1 \times 3$. We could also easily list all non-realizable integers.

3.2.4. *Example 2: Leaky fuel tank.*

> A car begins a trip with 40 liters of fuel and drives at a constant speed 70 km/hr. At this speed the car gets 12 km/liter, but it runs out of fuel after 360 km. Apparently the fuel tank is leaking. What is the rate (liter/hr) of the leak[4]?

3.2.5. *Model.* The rate of loss is (fuel Lost)/(Time); denote this by L/T. There are two rate equations, and the relation of (fuel Lost) to (fuel Used):

(1) (speed, km/hr)(Time, hr) = (distance, km)
(2) (distance rate, km/liter)(Used, liter) = (distance, km)
(3) (Lost, liter) = (total, liter) - (Used, liter)

Now put in the numbers from the problem statement:

(1) $70\,T = 360$
(2) $12\,U = 360$
(3) $L = 40 - U$

Solve to get $T = 36/7$, $U = 30$, $L = 10$, and therefore $L/T = 70/36 = 35/18 \simeq 1.94$ liters/hour.

[4]Adapted from another 'puzzler' from *Car Talk*. This has been translated from miles and gallons so I won't be embarrassed if someone from a modern country reads it.

3.2.6. *Discussion.* The first step is to write the rate equations in a "symbolic" form designed to be easily and reliably remembered. Neither numerical values nor *any* analytic reasoning are done in this step. Trying to look ahead, for instance "I need to know the time, so write the speed rate equation as (time)=(distance)/(speed)..." invites confusion. This problem is significantly more difficult without modeling because there are so many opportunities for such confusions.

Note also that the rate equations are written with units, so dimensional analysis can be used as a check that they are correct. In particular, the rate in the fuel use equation is given as a *distance* rate (km/liter) rather than a *fuel* rate (liter/km). The latter seems more logical (see the next problem), and confusion on this point is likely to be a source of error.

The final step is to substitute numbers into the symbolic forms, and then solve. This is routine, and easy because the analysis is protected from cognitive confusion with modeling.

3.2.7. *Example 3: Fuel efficiency.* This illustrates a common way to misuse statistics. Comparison with the degraded version illustrates the benefits of symbols rather than the usual inert numbers.

> A regulatory agency wants to promote development of fuel-efficient cars. The regulations require a certain *average* efficiency to provide automakers flexibility and incentive: they can offer inefficient models if these are balanced by super-efficient ones. However an environmental group claims that if the regulators use the traditional km/liter measure of automotive efficiency then more-efficient models will actually lead to *increased* fuel use. Are they right? Which measure should be used to avoid this?

3.2.8. *Symbolic specific problem.* Suppose there is a target efficiency T km/liter. A manufacturer sells two models: an efficient one that gets rT km/liter for some efficiency multiplier $r > 1$, and an inefficient one that gets b km/liter so that the average of rT and b is T.

(1) Find the average of the fuel required by the two models to go distance d. Express this as the product of the fuel required by a car with average efficiency T, and a function $h(r)$.
(2) Plot $h(r)$ (by calculator or computer) on the interval $[1, 1.8]$. For which efficiency multipliers r does the two-model fleet use more fuel than average-efficiency cars?
(3) Evaluate $h(r)$ for $r = 4/5$ and $5/3$. Find r so that the two-model fleet uses twice as much fuel as a fleet of average-efficiency cars.
(4) Explain what could theoretically happen if the automakers developed a car with efficiency *twice* the target (i.e. $r = 2$).

3.2.9. *Discussion.* The previous problem provides an appropriate template for the modeling step, so an explicit solution is not given here.

Better measure? Here we are interested in total fuel use. The km/liter measure has rate equation

$$\text{(fuel, liter)} = \frac{\text{(distance, km)}}{\text{(distance rate, km/liter)}}$$

so it is *inversely* related to fuel use. But standard statistical methods are additive: they work well if the data has a linear structure, poorly otherwise. This explains why the km/liter measure works poorly. The *linearly* related efficiency measure is the inverse, the fuel rate in liter/km with rate equation

$$(\text{fuel, liter}) = (\text{distance, km})(\text{fuel rate, liter/km})$$

3.2.10. *Degraded specific problem.* The regulatory agency sets 12 km/liter as the target efficiency. A manufacturer sells two models: an efficient one that gets 18 km/liter and an inefficient one that gets b km/liter so that the average of 18 and b is 12.

(1) Find the average of the fuel required by the two models to go 100 km. Compare to the fuel required by a car with average efficiency 12.
(2) Repeat for high-efficiency rate 20, and 22. Describe the pattern you see as the better efficiency increases.

The use of specific numbers is typical of contemporary practice. This makes it easier to do without modeling, and more accessible to direct evaluation on a calculator. On the other hand a great deal of mathematical structure has been hidden. The average fuel use as a function of the upper efficiency is now represented by a few numerical values. These suggest part of the pattern, but do not make obvious the infinite limit at upper efficiency 24. Further, the use of a specific numerical target hides the fact that the controlling parameter is (max efficiency)/(target efficiency), called r in the first version. All this structure emerges from essentially the same work done symbolically.

The use of numerical values is also a favorite tactic of test designers. The first version is a single problem. Plugging in numbers gives a large number of seemingly different problems. But don't we *want* students to see structure and patterns? Destroying it for the convenience of test design is not appropriate.

3.2.11. *Example 4: quarter-full fuel tank.* This[5] illustrates a difficulty with trick problems.

> A trucker's fuel gage is broken, and rather than fix it he puts a stick in the tank and measures the length that gets wet by the fuel. The tank is a circular cylinder with horizontal axis. He knows that when the wet length is half the diameter then he has a half tank of fuel. What length corresponds to a quarter tank of fuel?

This sounds like a typical geometry problem that can be solved by a clever trick or insight. One can waste a lot of time looking for the trick, but there isn't one. To *prove* there is no trick we find an equivalent problem. A nontrivial integration and trig identities show this is equivalent to the following:

- Find $s > 0$ so that $\cos(s) = s$
- define d by $(\frac{s}{2})^2 = d^2 - d^4$, then
- the length of wet stick corresponding to 1/4 tank is $1 - d$ (approximately 0.5960) times the radius of the tank.

The quadratic formula can be used to express d explicitly in terms of s using square roots, and conversely knowing d solves $\cos(s) = s$, so this has equivalent difficulty. But this is not an elementary problem.

[5] Yet another 'puzzler' from *Car Talk*.

The point is that trick problems are artificially contrived, and small changes give identical-sounding problems that are impossible with elementary methods. Trick questions give misleading impressions of the power of the methods and the nature of mathematics. This is a problem with Euclidean-style geometry, which is mostly tricks limited to special cases. For further discussion see §5.5.

3.2.12. *Example 5: crossing the river.* Another example paraphrased from [4]:

> A group of adults and children on a camping trip come to a river. They find a boat that can hold one adult or two children. Anyone in the group can safely row across the river by themselves. If there are four adults and two children on the trip, is it possible to get all of them across the river? If yes, how many one-way trips across the river will it take?

3.2.13. *Model.* We model this in terms of moves:

- A corresponds to an adult taking the boat from the first shore to the second; C, C^2 correspond to one or two children taking the boat. Denote the number of adults and children by a, c respectively.
- A sequence of crossings corresponds to a list (or "word") in which moves and their inverses alternate. For instance AA^{-1} corresponds to an adult taking the boat across and then coming back.
- A sequence is "allowable" if every initial segment has total exponent on A in $[0, a]$ and total exponent on C in $[0, c]$. For instance $AA^{-1}C^2C^{-1}$ is an allowable sequence if a, c are not too small, but $C^2A^{-1}C^{-1}A$ is not because the initial segment C^2A^{-1} has -1 total exponent on A.
- When is it possible to have an allowable sequence with total exponents a, c?
- What is the minimal length of such a sequence?

There are so few allowable sequences without immediate cancellations that the solution to the original problem can be found essentially by trying all possibilities. I describe the general solution.

First note that if $c \geq 2$ then C^2C^{-1} is an allowable sequence of even length with net effect C. Iterating this gives an allowable sequence with total exponent $c - 1$, and ends with C^{-1}. Omit this final move to get a sequence with exponent c. The total length of his sequence is $2c - 3$.

Returning to the word problem, if there are at least two children then all of the children can get across the river.

Now look for a similar even-length sequence with total effect A. It must start with C^2 or the next move will just cancel it. The second move must be C^{-1} for the same reason. The next move can be A. The fourth move must be C^{-1}: A^{-1} would defeat the purpose and C^{-2} is not allowable. We are thus led to $C^2C^{-1}AC^{-1}$. Iterating this gives an allowable sequence with total A exponent a, and C exponents between 0 and 2.

Combining the moves above shows: if $c \geq 2$ then any pair a, c can be realized, by a sequence of moves of length $4a + 2c - 3)$. This is the minimal length. Minimality can be proved by induction on a but this is not really elementary.

3.2.14. *Discussion.* This problem does not connect with much of anything. Most modifications of the rules, to have the boat carry one adult and one child, or up to three children, for instance, make the problem trivial. Variations with three passenger types, one of which might eat one of the others, are even more contrived and still fail to illustrate mathematical structure. Like so many trick problems, this seems to be a clever dead end.

3.3. **Integer partial fractions.** The objective here is to tie together the previous two sections, and foreshadow a standard rational-function topic in §3.5. This might also illustrate that there are lots of interesting problems within mathematics.

First, a theorem. Theorems are great labor-savers: they explain what you can do with complete confidence, and what the limits are. In this case, for instance, you should expect that attempts to separate non-relatively-prime factors will fail in some way.

Theorem 1. *If b and c are relatively prime, then fractions with denominator bc can be expanded as $\frac{a}{bc} = \frac{x}{b} + \frac{y}{c}$, for some x, y.*

This is true essentially whenever the terms make sense (any commutative ring) and is an immediate consequence of the definition of relatively prime. a, b are *relatively prime* if there are m, n so that $am + bn = 1$. This is the correct general definition but it is cumbersome to use, and in practice (integers here, and real polynomials in §3.5) one uses refinements of both 'relatively prime' and the expansion.

Theorem 2. (Refinements for integers):

(1) *For integers, 'relatively prime' is the same as 'have no common factor' (except 1, of course).*
(2) *if $b_1, b_2 \dots b_n$ are pairwise relatively prime then there is a unique expansion*
$$\frac{a}{b_1 \times \cdots \times b_n} = \frac{r_1}{b_1} + \frac{r_2}{b_2} + \cdots + \frac{r_n}{b_n} + q$$
with all the $\frac{r_i}{b_i}$ proper fractions and q an integer.

For integer fractions, $\frac{a}{b}$ *proper* means that $b > a \geq 0$; see §3.5 for the polynomial version. Note that the version for a single factor, $\frac{a}{b} = \frac{r}{b} + q$ with $\frac{r}{b}$ proper, is the usual quotient-with-remainder expression for $a \div b$.

This refinement is true because the Euclidean algorithm works in the integers. In fact there is an algorithm for the numerators in the expansion, based on the extended Euclidean algorithm (see the Wikipedia entry). The algorithm is cumbersome and a bit obscure, so a modular-arithmetic approach is used here.

3.3.1. *Problem.* Express 47/180 as a sum of an integer and proper fractions with prime-power denominators.

3.3.2. *Solution.* $180 = 2^2 3^2 5$ so the partial fraction expansion is of the form $\frac{a}{4} + \frac{b}{9} + \frac{c}{5} + d$.

The first step must be to clear denominators: fractions are defined implicitly, and we must convert to an explicit form to work with them (see §3.1). This gives
$$47 = (4 \cdot 9 \cdot 5)(\frac{a}{4} + \frac{b}{9} + \frac{c}{5} + d) = 9 \cdot 5a + 4 \cdot 5b + 4 \cdot 9c + 4 \cdot 9 \cdot 5d.$$

Reduce mod 4 to get $47 \overset{4}{\equiv} 3 \overset{4}{\equiv} 45a \overset{4}{\equiv} a$. This means $a = 3 + 4r$, and the numerator that gives a proper fraction is 3. Next reduce mod 9 to get $2 \overset{9}{\equiv} 2b$, so $b \overset{9}{\equiv} 1$. Finally reduce mod 5 to get $2 \overset{5}{\equiv} c$.

We now know that $47/180 = \frac{3}{4} + \frac{1}{9} + \frac{2}{5} + d$, with d an integer. We can find d by converting the partial fractions back to a common denominator, but we could also use a calculator. The left side is approximately 0.2611, while the fractions on the right add up to about 1.2611. Since d is an integer it must be -1, and

$$47/180 = \frac{3}{4} + \frac{1}{9} + \frac{2}{5} - 1$$

3.3.3. *Back to chicken nuggets.* We could reformulate the mathematical model of the chicken-nugget problem as: For which a is there a partial fraction expansion

$$\frac{a}{180} = \frac{\square}{30} + \frac{\square}{20} + \frac{\square}{9}$$

with *nonnegative* numerators? Or equivalently, as a sum of proper fractions with denominators 30, 20, and 9, and a *nonnegative* integer?

As mentioned in the discussion of the problem, there is always an expansion if negative numerators are allowed. This situation is not covered by Theorem 1 because the factors are not pairwise relatively prime, but thinking about this case clarifies the role of the hypotheses in the theorem. There is always an expansion

$$\frac{a}{b} = \frac{r_1}{b_1} + \frac{r_2}{b_2} + \cdots + \frac{r_n}{b_n} + q$$

if every prime power divisor of b also divides one of the b_i. The numbers in the chicken-nugget problem satisfy this condition. The expansion is unique if the b_i are relatively prime. This is not true in the problem, and examples are easy to find. The existence and uniqueness observations together give the theorem because if b is the product of the b_i then b_i relatively prime implies the prime-power divisor condition.

In the opposite direction, the denominator-cleared form of the 47/180 problem above is: obtain 47 by adding multiples of 45, 20 and 27. The appearance of the -1 in the answer shows that it cannot be done with nonnegative coefficients.

3.4. **Multiplying polynomials.** This section illustrates how careful attention to cognitive issues and mathematical structure can significantly extend the problem types students can do. The main cognitive concern is that different activities (here organization, addition, and multiplication) should be separated as much as possible. Structure is used to make the procedure flexible; see the next section for a variation. See [7] for detailed discussion and extensions.

3.4.1. *Problem.* Write the product $(2y^3 + y^2 - 9)(-y^2 + 5y - 1)$ as a polynomial in y, in standard form.

Here "standard form" means coefficients times powers of the variable, with exponents in descending order. The factors are given in this standard form. Sometimes ascending order is used.

3.4.2. *Solution.* The first step is purely organizational. We see that the outcome will be a polynomial of degree 5 so set up a template for this:

$$y^5(\qquad\qquad\qquad) + y^4(\qquad\qquad\qquad) + y^3(\qquad\qquad\qquad) + \cdots$$

Next scan through the factors and put coefficient products in the right places. To get the y^3 coefficient, for example, begin with the leftmost term in the first factor. Record the coefficient, (2). The exponent is 3, so the complementary exponent is 0. Beginning at the right in the second factor, we see that the coefficient on y^0 is

-1, so we record this as a product $(2)(-1)$. Now move one place to the right in the first factor and record coefficient (1). The exponent is 2 so look for complementary exponent 1 in the second factor. The coefficient on this is 5, so record this as $(1)(5)$. Again move one place to the right in the first factor and record coefficient (-9). The exponent is 0 so we look for complementary exponent 3 in the second factor. There is none, so we record coefficient 0 as $(-9)(0)$. The template will now look like

$$\cdots + y^3\big((2)(-1) + (1)(5) + (-9)(0)\big)\cdots$$

It is quite important that *no arithmetic* be done in the organizational phase, not even $(1)(5) = 5$. Also, we usually quit scanning the first factor when complementary exponents are above the degree of the second, but it is better to record overruns such as $(-9)(0)$ than to worry about this. The reason is that even trivial on-the-fly arithmetic requires a momentary break in focus that significantly increases error rates in both organization and arithmetic. For the same reason, every coefficient is automatically enclosed in parentheses whether they are needed or not.

The first arithmetic step is to do the multiplications:

$$y^5\big(\underbrace{(2)(-1)}_{-2}\big) + y^4\big(\underbrace{(2)(5)}_{10} + \underbrace{(1)(-1)}_{-1}\big) + y^3\big(\underbrace{(2)(-1)}_{-2} + \underbrace{(1)(5)}_{5} + \underbrace{(-9)(0)}_{0}\big) + \cdots$$

The second arithmetic step is to do additions:

$$y^5\big(\underbrace{(2)(-1)}_{-2}\big) + y^4\big(\underbrace{\underbrace{(2)(5)}_{10} + \underbrace{(1)(-1)}_{-1}}_{9}\big) + y^3\big(\underbrace{\underbrace{(2)(-1)}_{-2} + \underbrace{(1)(5)}_{5} + \underbrace{(-9)(0)}_{0}}_{3}\big) + \cdots$$

Additions and multiplications are separated because, again, switching back and forth invites errors. Note also that this pattern requires essentially no organizational activity during the arithmetic steps: everything is in standard places, and outcomes are written just below inputs.

The final result is

$$y^5(-2) + y^4(9) + y^3(3) + y^2(8) + y^1(-45) + 9$$

3.4.3. *Discussion.* Standard high-school practice is to restrict to product of binomials, and use the intelligence-free algorithm with acronym “FOIL”. This is so ingrained that many students say “FOIL it out” rather than “expand”, even when the factors are not binomials. However this algorithm mixes organization and arithmetic enough that some students have trouble even with this simple case, and since the expansion step is not adapted to the eventual goal, a second sorting step is needed. Finally, the near-exclusive focus on binomials makes larger terms foreign territory. Some students are at a loss about how to deal with them, others generalize incorrectly from the pattern, and almost all are anxious and hesitant. A better algorithm fixes all this, and opens up a much wider world.

3.5. **Polynomial partial fractions.** The final example brings together fractions, partial fractions, polynomial multiplication, and even modular arithmetic. Theorem (1) in §3.3 applies, but as in that section we need a refinement.

Theorem 3. (Refinement for polynomials):

(1) *For polynomials ‘relatively prime’ is the same as ‘have no common factor’, and the same as ‘have no common roots’ (including complex roots).*

(2) *If $b_1, b_2 \dots b_n$ are pairwise relatively prime real-coefficient polynomials then there is a unique expansion*

$$\frac{a}{b_1 \times \cdots \times b_n} = \frac{r_1}{b_1} + \frac{r_2}{b_2} + \cdots + \frac{r_n}{b_n} + q$$

with r_i, q polynomials and each $\frac{r_i}{b_i}$ a proper fraction.

A *polynomial* fraction is called proper if the degree of the numerator is strictly smaller than the degree of the denominator. The term has the same significance as the integer version even though the definitions are different. As with integers, q is the result of division (the "quotient") and the r_i are remainders.

This is statement almost identical to the integer statement. This emphasizes commonality of the underlying structure, but the main reason is efficiency: since the underlying structures *are* the same, a maximally efficient statement for one context will also be maximally efficient for the other.

3.5.1. *Problem.* Find the real partial-fraction expansion of the polynomial fraction

$$\frac{3x^3 + 2x + 9}{(4x^2 - 4x + 1)(x^2 - 2x + 3)}$$

3.5.2. *Solution, setup.* The first quadratic in the denominator factors as $(2x-1)^2$, but these factors cannot be separated because they are not relatively prime. The second quadratic has complex roots so it can be factored over the complexes as a product of linear terms. These could be separated in a partial-fraction expansion over $\mathbb{C}$, but not over $\mathbb{R}$. The denominators in the partial fractions are therefore the two quadratics. Finally, the input fraction is proper so the expansion cannot have a (non-fraction) polynomial term (note this conclusion is special to polynomials: it does not work for integers).

Both pieces in the expansion have denominators of degree 2, and we know that the numerators have smaller degree. The expansion is therefore of the form

$$\frac{3x^3 + 2x + 9}{(4x^2 - 4x + 1)(x^2 - 2x + 3)} = \frac{ax + b}{4x^2 - 4x + 1} + \frac{cx + d}{x^2 - 2x + 3}$$

As usual, to work with fractions we have to clear denominators. This gives

$$3x^3 + 2x + 9 = (ax + b)(x^2 - 2x + 3) + (cx + d)(4x^2 - 4x + 1) \tag{1}$$

I describe two ways to find the unspecified coefficients.

3.5.3. *Solution, linear-algebra approach.* Equation (1) is an equality of polynomials, so the coefficients must be equal. The coefficient equations give the linear system that determines a–d.

The plan is as follows: for each exponent n scan through the products above and pick out coefficients on x^n, exactly as in the section on polynomial products. Recording coefficients gives

$$\begin{array}{ll}
x^3: & 3 = a(1) + c(4) \\
x^2: & 0 = a(-2) + b(1) + c(-4) + d(4) \\
x^1: & 2 = a(3) + b(-2) + c(1) + d(-4) \\
x^0: & 9 = b(3) + d(1)
\end{array}$$

Writing this linear system in matrix form gives

$$\begin{pmatrix} 1 & 0 & 4 & 0 \\ -2 & 1 & -4 & 4 \\ 3 & -2 & 1 & -4 \\ 0 & 3 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 2 \\ 9 \end{pmatrix}$$

For this approach a good problem statement is "find a linear system that determines the coefficients…", so this would the solution to the problem.

If a full solution is required then the problem statement should be "find a linear system that determines the coefficients, and then solve to find the expansion". Requiring explicit display of the intermediate step has two functions: first, if the final answer is wrong then the intermediate step can be used to locate errors. The second reason is to ensure that students carefully formulate the intermediate step, especially if the system is to be solved by hand. Many will try to save writing (and thinking) by solving on the fly as coefficients are found (e.g. for the x^3 coefficient writing $a = 3 - 4c$ and using this to eliminate a). They could use this strategy to solve the system after it is set up, but mixing the steps increases the error rate and in the long term will limit the problems they can handle. In the long run they would enter

```
LinearSolve[{{1,0,4,0},{-2,1,-4,4}, {3,-2,1,-4},{0,3,0,1}},{3,0,2,9}]
```

in a computer-algebra system[6], and obtain $(\frac{319}{81}, \frac{214}{81}, -\frac{19}{81}, \frac{29}{27})$.

As a final note, if the input fraction were not proper then the q term in Theorem 3 would be nonzero. It can be included in the cleared-denominator form (1) and handled the same way. Most approaches recommend first using long division to get q (the quotient) and a proper fraction (with remainder as numerator), and then expanding the proper fraction. The division algorithm gives the coefficients in the quotient, so the remaining coefficients give a smaller linear system. But division takes a lot more work. The extra part of the larger linear system has lower triangular coefficient matrix so is quick and easy to solve.

3.5.4. *Solution, modular arithmetic approach.* In the integer version there is no analog of the linear algebra approach, but modular arithmetic is successful. There is a polynomial analog of modular arithmetic that I illustrate by finding the coefficients c, d in equation (1).

The plan is that we want to work modulo the polynomial factor on the other term, $x^2 - 2x + 3$, so set $x^2 - 2x + 3 \equiv 0$. Technically we are working in the quotient ring $\mathbb{R}[x]/(x^2 - 2x + 3)$. In practice we write the imposed identity as $x^2 \equiv 2x - 3$ and use this to reduce second and higher-order terms. For instance $x^3 = x(x^2) \equiv x(2x - 3) = 2x^2 - 3x \equiv 2(2x - 3) - 3x = x - 6$. Equation (1) reduces to

$$5x - 9 \equiv (cx + d)(2x - 3).$$

Expand the right side and reduce the second-order term to get $5x - 9 \equiv (-3c + 4d)x + (-12c - 11d)$.

We now apply the fact that if two degree-one polynomials are equivalent modulo a degree-2 polynomial then they must actually be equal. (The technical context is

[6]This is *Mathematica* syntax.

that this is a Gröbner basis for the polynomial quotient ring.) The coefficients on the two sides must therefore be equal and we get the system of equations

$$\begin{pmatrix} -3 & 4 \\ -12 & -11 \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} 5 \\ -9 \end{pmatrix}$$

This implies $(c, d) = (-\frac{19}{81} + \frac{29}{27})$, as before. Clearly, the other term can be done the same way.

A version of this is commonly used as a shortcut: working modulo a degree-one polynomial $x - r$ is the same as evaluating at $x := r$. If the factors of the denominator are *distinct* linear terms (i.e. no complex or repeated roots), and if it is easy to do arithmetic with the roots (i.e. rational), then the coefficients can be obtained by evaluations. Some high-school courses use this method exclusively. This is a bad idea, in the same way that restricting to quadratics with integer roots is a bad idea: it enables easy methods that don't work in more general (and realistic) cases.

In general the modular-arithmetic version seems to take longer than linear algebra. It is worth exploring to illustrate similarities between the integer and polynomial situations. It is also good to give students multiple tools, and sometimes hybrids can be used to good effect. For example, if there are n unknown coefficients then the goal is to find n linear equations that determine them. The routine system comes from coefficients in the polynomial equation. However if there is a root at which polynomials can be easily evaluated (e.g. small integer) then the evaluation can be used to get one linear equation that is often simpler than the ones from coefficients. Replace one of the coefficient equations with this to get an easier system. This comes with a caution: the system from polynomial coefficients is guaranteed to be nondegenerate, so can *always* be solved. Replacing one of the equations may give a degenerate system. It is a good strategy because the failures are very rare, but students should know how to recognize a failure and what to do about it (go back to the routine system).

4. Mathematics and teachers

As explained in the introduction, the primary output from the science-of-learning approach is a toolkit of techniques, definitions, etc. designed for student use. This does not come with assembly instructions: how to put it all together to get courses and a curriculum is the job of teachers. Getting the components to actually work, however, puts strong constraints on methodology. Implications for methodology are described in the second subsection. The first provides a convenient context. The final subsection draws conclusions about teacher preparation.

4.1. **Mathematics.** This is an overview of structure and practice of mathematics, formulated to suggest use by young learners, and to guide development of material for young learners. These are extracts from a description of mathematics developed bottom-up from details and concrete observations, [6], but what makes it valuable is that it accurately encodes and organizes methodological constraints of the micro-scale material here and in [7]. This description has little in common with the hypothetical constructs of philosophers.

4.1.1. *Reliability.* Mathematics has evolved a set of explicit rules of reasoning with the following astounding property: arguments without rule violations have completely reliable conclusions[7]! People have trouble learning to use these rules effectively because rules in most other systems are sufficiently vague and ineffective that precision is a waste of time, while in mathematics anything less than *full* precision (*no* rule violations) is a waste of time. However mathematical rules are simple compared to those of the tax code, religions, politics, physical science, etc. Most people can master the basics, and within its domain of applicability this can be a powerful enabling technology.

4.1.2. *Potential proofs.* For users the key enabling concept ([6] §?) is *potential proof*: a record of reasoning, using methods of established reliability, and detailed enough to check for errors. Potential proofs in this sense are very common. The scratch work done by a child multiplying multi-digit integers by hand provides a record of reasoning that can be checked for errors. When teachers tell students to "show your work" they mean "provide a record of reasoning…". The basic method is already in wide use, and the potential-proof description mainly clarifies the features and activities needed to make it fully effective.

In these terms a "real" proof is a potential proof that has been checked (and repaired if necessary) and found to be error-free. However the benefits come from reasoning and checking. Errors happen. A flawed first attempt with sufficient detail can provide a framework for diagnosis and improvement. A flawed first attempt with insufficient detail is useless and the user must start again from scratch. Focusing on the formal and error-free aspects of completed proofs therefore distracts from the features that actually provide power to users.

4.1.3. *Definitions.* The next point is that precise reasoning about something is only possible when the thing is described precisely. Vague or intuitive things are inaccessible to mathematics. This is not to say intuition is irrelevant, but effective mathematicians do not work directly with intuitions. They make a precise description that they hope will realize their intuition, work with it for a while, and if it doesn't do what they want then they junk it and try again with a different precise description. In some areas these false starts probably account for 50% of all mathematical effort, and when a really good precise description is finally found it is a prize and a treasure. The idea that mathematical definitions are random is as nonsensical as the idea that mathematics can be done with intuitions.

Concise modern definitions provide the precision needed for mathematical reasoning, but the format itself is not forced by the job to be done. Details of the format seem to have evolved to be effective for human use; see [6] §?. In other words, precise definitions are another enabling technology for people, not unnatural constructs.

4.1.4. *Limits of mathematics.* The final point is that the need for complete precision limits the scope of mathematics. In particular, nothing in the physical world can be described with mathematical precision, so mathematical methods do not apply directly. People applying mathematics accommodate this with an intermediate step: a symbolic *mathematical model* is developed to represent the physical situation, and it is the model that is analyzed mathematically. Mathematical conclusions

[7] For all practical purposes; see [6] §? for nuances.

about the model are reliable in a way that the connection between the physical situation and the model cannot be, and working without a model confuses this structural difference. In Section 3.2 I emphasized the importance of modeling to avoid cognitive interference. Cognitive and structural differences should probably be seen as *different* reasons to separate the two activities, rather than cognitive interference seen as resulting from structural difference. There are plenty of things (e.g. multiplication and addition) that to some degree interfere cognitively but are not deeply different structurally. Or put more directly, brains are strange things and it is a serious mistake to imagine that brain structure reflects the structures we imagine in the world.

4.2. **Teaching.** These suggestions summarize methodology needed to make the examples like those in §3 work effectively. However they should be thought of as ways to avoid known modes of failure rather than as certain paths to success. This reflects the scientific approach: success in science is not direct, but a matter of recognizing and learning to avoid ways to fail.

These suggestions are strongly at odds with mainstream educational methodology. Put another way, there are many modes of failure that mainstream methods do not avoid. This point is sharpened in the section on teacher preparation.

4.2.1. *Precision.* Precision makes mathematical work go smoothly, and this is particularly true in elementary mathematics. Sloppy thinking that is harmless to experienced users can seriously confuse beginners. Section §5.1 describes confusion about fractions resulting from failing to distinguish between things and names for them; between representatives of an equivalence class and the equivalence class; and from common ambiguity in the use of '='. The following subsections suggest ways to avoid these and other problems.

4.2.2. *Names.* Distinguishing between things and names for things may seem far too sophisticated for second-graders, and it would be a mistake to formulate this as an abstract idea. However simply *doing* it as a matter of routine could clarify many things:

- Fractions and decimals are different names, not different things (§3.1). Trying to make sense of them as different things cannot be successful.
- Expressions like $34.1/5.2$ and $7.7 + 22.2$ specify numbers. Performing the indicated operations is not "finding the numbers" but expressing them as decimals, i.e. finding the decimal names.
- We are on a first-name basis with the integers $0 \dots 10$ and a few others (π, e). Everyone else has a compound name. The Babylonians used a base-60 system, which meant they had to be on a first-name basis with the first 60 integers! This was way too many for easy use.
- The Roman numeral system has a lot of numbers known by first names: 1=I, 5=V, 10=X, 50=L, 100=C, 500=D, 1000=M,.... This is awkward and is one reason we don't use it anymore.
- "John + Alice" is carved on a tree. If we decide "John" is a name for 5, and "Alice" is a name for 7, what is the number that corresponds to the carving?

There are further possibilities. Customary names for numbers often do not reflect mathematical structure. 13, for instance, is a compound name but "thirteen" is a cognitive unit. Cross-cultural neuroscience studies suggest that the cognitive overhead associated with customary names makes arithmetic more difficult. Use of "math names" (a sequence of digits; "one three" for 13 for example) might therefore make arithmetic easier. See *Neuroscience experiments for mathematics education* in [7] for more detail.

Finally, *indirect* names encode a property that characterizes something. Being more explicit about the nature of indirect names would clarify the standard examples, including fractions and roots. This might also help liberate us from obsolete notational constraints. For example "the largest real root of $5x^5 - 3x^3 + 2x + 1$" specifies a number just as well as $\sqrt{37}$ does, and these days it can be numerically approximated just as easily. Currently we focus on quadratic polynomials because the quadratic formula provides names for the roots in terms of square roots. But this is a strong constraint and the resulting mindset ("quintics are impossible") is a liability in the long run. Is it time to outgrow this?

4.2.3. *Definitions.* Important objects and properties (fraction, prime, relatively prime, proper fraction, . . .) should be given brief and genuinely precise definitions.

- These definitions should be constructed primarily by professional mathematicians. Professionals hone and fine-tune definitions to be brief and effective, and will ensure that they are compatible with later material.
- Students should be required to memorize them so they can be reproduced exactly. Definitions provide anchor points, and understanding nucleates and deepens around the definition. This is particularly true for weaker students.
- Explanations of what a definition 'means' should be given after the definition, not before. Putting the explanation first almost guarantees confusion.

Compared to current educational practice this is very rigid, but it is essentially the way mathematicians approach unfamiliar material, and it works. More precisely, less rigid approaches take longer, are less effective, and were abandoned when this approach became available about a century ago.

Is this approach reasonable in education? Well-crafted definitions evoke mathematical objects with the economy and grace of poetry. Asking students to memorize them is like asking them to memorize unusually powerful poems. There are not so many that this is unreasonable, and if it is done consistently it will become routine. Students will also see quick payoffs because good definitions are immediately functional.

Current practice avoids concise definitions, ostensibly to focus on "understanding". Students may be invited to put the idea in their own words or even make up their own version. This seems to be a carryover from philosophy. If one thinks of mathematical things as tools, then it makes sense to provide students with the sharpest versions and focus on using them. Philosophical ideas are not particularly functional as tools, and in that tradition the emphasis is on "understanding" and appreciating ideas rather than actually trying to use them.

In a mathematical approach students might still be invited to rephrase definitions once they can do so with precision, but if they get it wrong they *must* be told so. To do otherwise deprives them of a solid foundation and will cause trouble later. Corollaries are that students should be invited to reformulate definitions *only* if the teacher has the time and expertise to diagnose the reformulations for correctness, and *only* if it seems likely that students can understand, accept, and benefit from corrections.

4.2.4. *Word problems.* Word problems have two components: the word part, and the mathematical core. Serious users employ mathematical models as an intermediate step to separate these activities. Contemporary educators, in contrast, are philosophically committed to word problems as a different *format* rather than a different activity. They encourage a holistic approach with reasoning 'in context', and discourage modeling.

Historically, systematic use of modeling developed in the seventeenth century, and was an essential part of the 'scientific revolution'. The contemporary educational approach follows sixteenth century practice. But this has been obsolete for four centuries for good reasons: the two components are different both cognitively (different brain regions) and structurally. Mixing them confuses and degrades both. In fact cognitive interference in word problems may be the most serious single difficulty I have seen in students. For weaker students "crippling" may not be too strong a word.

Difficulties currently caused by word problems are unnecessary: modeling can be used easily and effectively in school mathematics. Examples are given here in §3.2, and there are many more in [7].

Summary: in recent years mathematics educators have increased emphasis on word problems, but insist on using an obsolete approach that makes them unnecessarily difficult and inhibits rather than promotes mathematical development. Fixing this would probably be the most effective single step that could be taken to improve K-12 mathematics.

4.2.5. *Technology.* Modern technology is powerful, and a great deal of [7] is concerned with educational applications. The issue is too complex to be usefully addressed addressed here, but I make one comment.

The traditional classroom was a tightly bound package, and technology is making it come apart. In particular it is loosening the link between teaching and learning. For instance calculators make teaching easier, and improve student performance on traditional measures (test problem types). However these measures no longer have the learning correlates they once did. Better teaching and test performance are, in some cases, masking learning declines. A genuinely learning-oriented approach is essential for good long-term outcomes.

4.2.6. *Work templates.* Students should be taught to record their reasoning in a way that can be checked for errors (a "potential proof" in the terminology of §4.1.2). What to do with such a record is the topic of the next section; here the concern is with how to get them to do it.

Students will not use formats that are cumbersome or require a lot of extra writing. Formats that are complicated or distracting are counterproductive. The criteria for good work formats are therefore:

(1) record enough detail so reasoning can be reconstructed and checked for errors;
(2) be compact and straightforward; and
(3) help organize the work in ways consistent with human cognitive constraints.

Making up such a format is hard for professionals and cannot be expected of students. Students must be provided with *templates* carefully designed to meet the criteria above.

The standard formats for multi-digit multiplication and long division are well-known examples of templates, though they are not fully satisfactory examples. They satisfy the first two conditions, but are less successful with (3) because they have been optimized for production arithemetic by experienced users rather than to provide support for beginners. A less-efficient but more supportive template for multiplication is explored in *Neuroscience experiments for mathematics education* in [7]. This template for numbers is largely derived from a template for multiplying polynomials described in §3.4.2 and elsewhere.

The underbrace notation in §3.4.2 for processing parts of a complex expression is a particularly useful template. Students do have to be shown how to carefully collect the processed fragments, however.

4.2.7. *Reading and writing complex expressions.* Working with mathematical expressions requires that they be parsed (linearized), and there are a number of ways to do this. The logical structure encodes an "outside-in" parsing order. For example:

$$5(3y^2 - a) + (y - a)(y + 6a)$$

is the sum of two terms, and each of these terms is a product. Logically parsing to this level gives

$$5(\dots) + (\dots)(\dots).$$

Terms in parentheses are deeper in the logical structure. In contrast the left-to-right parsing used for reading gives "five times the quantity three wye squared dots". The encapsulation by parentheses is hard to say and may get lost. There is also an "inside-out" parsing order reflecting the fact that processing usually begins with innermost fragments and works out. This leads many students to ignore large-scale structure until inner things are simplified. For instance they will want to expand the $(y - a)(y + 6a)$ term before even thinking about the rest of the expression. Teaching students to use logical (outside in) parsing order extends the expressions they can process, but the greatest benefits are in the expressions they can write. The organizational step in the polynomial multiplication template §3.4, for instance, uses mathematical (outside in) parsing very explicitly by writing parentheses first and filling them in later. Another instance is the summation notation. If this is not parsed correctly it is hard to make sense of it, and almost impossible to write or manipulate it correctly.

I have two concrete suggestions. First, the logical structure can be emphasized when expressions are described by teachers. For example when writing the expression above, after writing "5(" write the corresponding *closing* parenthesis first, and then fill in. Students emulate what they see. Note that this requires a dynamic presentation: showing the end result does not show this structure. A corollary is that to be maximally effective, texts and reference materials really should show the process, not just the outcome (YouTube?!).

The other concrete suggestion has to do with parentheses. The customary notation is hard to parse, and elementary educators go to great lengths to avoid them (indeed "simplify" is very close to "write without parentheses"). But this makes complex expressions impossible and simple ones harder (see §3.4 where profligate use of parentheses is used to separate organization and arithmetic). Teachers often change the order of terms and do arithmetic on the fly to avoid intermediate forms with parentheses. These hidden operations often mystify students, and if they try to emulate this but come out with a form that is incorrect without parentheses then they get it wrong. The suggestion: use parentheses from the very beginning of elementary mathematics, with a notation that is easy to parse. Paired parentheses could be joined by an overline, for example:

$$5\left(\overline{3y^2 - a}\right) + 3\left(\overline{\left(\overline{y - a}\right)\left(\overline{y + 6a}\right)}\right)$$

4.2.8. *Accuracy and diagnosis.* The unique feature of mathematics is that it is possible to achieve near 100% accuracy. Extended arguments depend on this. If a single operation can be done with 70% accuracy then the likelihood of getting the right answer in problems with ten operations is under 3%. If one wants 70% accuracy in the ten-operation problem then individual operations must be done with 97% accuracy. The point is that elementary teachers who accept 70% success rates are not only wasting the unique potential of mathematics, but setting their students up for failure in later courses.

The proposal is that the goal in math courses should be quality, not quantity and not speed. Fewer problems, but expected to be 100% right. Is this reasonable in education? Maybe not, but I sketch strategies that should make it possible to get close.

First, students should be shown carefully-designed worked examples that they can emulate. Sometimes it is useful to have two versions: one with reasoning made explicit, and another that illustrates use of an efficient template or procedure in which reasoning is not explicit but can be reconstructed from the written record. Students who can omit minor detail without making errors can be allowed to do so, with two provisos. One is that they should still be required to use templates when these are designed to work for use with later problem types. For instance *Neuroscience experiments for mathematics education* in [7] gives an algorithm for multiplying multi-digit integers based on the polynomial algorithm in §3.4 and using the same templates. Familiarity with the templates is one of the objectives. The second proviso is that they should be required to fully record intermediate outcomes between cognitively different steps. Examples are models in word problems, and the intermediate between organization and arithmetic in polynomial multiplication. This is a vital work habit and allowing students to skip it will cause trouble later. Students who don't think they need to do this should perhaps be given more complicated problems.

The third strategy is teacher diagnosis of errors that students can't find themselves. Ideally *every* wrong answer should corrected, either by the student or through teacher diagnosis. For diagnosis the student should explain his reasoning, following his work record. If the record is illegible, steps were skipped, or appropriate templates were not used, then the work should be redone before being diagnosed for specific errors: it is remarkable how often this resolves the problem, and it is valuable for students to see this. If the student has an appropriate work record then it can be reviewed efficiently and mistakes quickly pinpointed. The teacher should wait until the student has come to the error before doing anything, however. Premature guesses about the difficulty are often wrong and will make confusion worse. Further, the student learns more if he spots the error himself; wait and see if he says, "oh, now I see what went wrong". In such cases the teacher can ask what happened, to see if the insight is right, but if so then the teacher should leave well enough alone and *not* explain further, even if a clearer description can be given. If this approach is used systematically then errors will become remarkably uncommon, because confusions don't accumulate and students learn to diagnose their own work. When teacher diagnosis is necessary it is fast and efficient.

Students unused to diagnosis may say "can you show me how to do this?" or "I don't even know how to start." The teacher should say "give it a try, and let's see how it goes", *not* "I'll show you". My experience is that are more open to this if the person doing the diagnosis is not also the person who will assign a grade.

A final strategy in this approach is the occasional use of long problems. Careful work with extensible methods should enable students to do much more complicated problems than the ones generally assigned. They should occasionally be given such problems, possibly as group exercises. There are two goals: first to stress their work habits a bit to see if they are solid, and second to show them that—with good work habits—long problems are really just more of the same, and their methods are considerably more effective than the focus on short problems might suggest.

I close the section with a caution related to human learning. Emphasizing quality over quantity—and carefully monitoring quality—should mean that boring repetition can be avoided: students learn at different rates, and ones who have learned the material can progress to more elaborate problems or the next topic. *However*, this does not mean students can stop as soon as they can use the methods accurately. Learning fades if practice is stopped too soon; durable learning requires practice or periodic reenforcement well past achievement of accuracy[8]. This is a delicate balance and needs to be carefully explored.

4.3. **Teacher preparation.** Outcomes from K-12 education are unsatisfactory and not improving. Angst about this has mainly focused on teacher preparation, implicitly assuming that educational methodology is effective and the problem is incompetent teachers. The discussion here suggests an alternative: teachers are fine but the methodology is incompetent, and worse outcomes might actually reflect *better* teaching of the methodology. If this is the case then the methodology has to be straightened out before there is much point in discussing teacher preparation.

The issue is approached through analysis of two common but conflicting viewpoints about advanced study.

4.3.1. *Advanced study* ***not*** *necessary?* The school-of-education argument is roughly "our students have trouble with advanced courses and don't get much from them. Pedagogy is more important than content anyway."

Wu [9] tries to give this point of view more substance by citing an article of Begle that found no correlation between advanced coursework and better student outcomes. Wu concludes that advanced viewpoints are irrelevant to elementary education, and uses fractions to illustrate the reasons. However what his example actually shows (see §5.1 below) is that imprecision and sloppy notation that are harmless at advanced levels can render the presentation irrelevant to elementary teaching.

[8]Cognitive psychologists refer to durably-learned material as "overlearned". This term is unfortunate because people interpret it as "excessive".

I would agree that courses such as real and complex analysis are unnecessary: they have little to do with K-12 topics and benefits to survivors is unlikely to justify the severe attrition among potential teachers. Courses such as abstract algebra could be more relevant. See §3.1 for a student-oriented presentation of fractions, and *Proof projects for teachers* in [7] for an extended development done relatively painlessly in commutative rings. This includes dealing with zero-divisors in general rings; not K-12 classroom material but it certainly clarifies why having 0 in the denominator is a bad idea. It also connects with modern themes by presenting Grothendieck groups as just a fancy name for fraction-type constructions. Another nice project for a course for teachers is to use modular arithmetic to *prove* the existence and uniqueness of the partial-fraction expansion described in §3.3.

These examples suggest that advanced courses with appropriate topics and precision could connect nicely with K-12 *topics*, but the next discussion reveals a problem with *methodology*.

4.3.2. *Advanced study **is** necessary?* Nearly all people with technical backgrounds feel teachers need more mathematics, but cannot clearly articulate why. It seems to me the essence is "my background leads me to feel that educators are doing bad things. If they had stronger backgrounds they would feel the same way, and do better." Roughly, the need is for more sophistication and mathematically disciplined thinking. The standard way to get these is subliminally from the material during extensive study, so more study seems to be the solution.

According to Beale, however, the teachers who did have enough advanced study to internalize the mindset don't do any better as teachers. In fact quite a few accomplished mathematicians have gone into education and most of them end up doing things that make technically-oriented people cringe.

There is a disturbing explanation for this. Advanced study leads to internalization of *modern* methodology. Education is modeled on a philosophical description of mathematical practice in the nineteenth century and before[9], centered around romanticized ideas about the power of intuition. In professional practice this was abandoned as ineffective in the early twentieth century. However educators reject contemporary methods as inappropriate for normal humans (young ones anyway). The bias against modern methods makes internalization from advanced courses useless. In fact the conviction that the nineteenth century was a better time for children is so strong that mathematicians working in education tend to buy into it, and suspend their professional skills and methods.

4.3.3. *Conclusions.* Genuinely productive discussion of teacher preparation must wait on resolution of deeper problems:

(1) Educational methodology for mathematics is stuck in the nineteenth century.
(2) Until this changes, we cannot expect outcomes better than those of the nineteenth century.
(3) Until this changes, advanced coursework with twentieth-century methodology will be irrelevant.

On a brighter note, the science-of-learning methodology described above is compatible with modern mathematics, and advanced coursework would be relevant to teaching based on this approach.

[9]For extensive historical and technical background on this see [6].

5. Comparisons

In this section I discuss some of the articles in the *Notices* special issue. The short version is that they are all seriously flawed, but to be useful (or even credible) this needs to be explained in detail. I focus on articles with enough explicit detail to support analysis.

5.1. **H. Wu.** Wu, in his article *The Mis-Education of Mathematics Teachers* [9], observes that there is a chasm between advanced coursework in mathematics and what is needed for contemporary K-12 teaching. He sees this as a problem with advanced coursework rather than contemporary teaching.

First, Wu points out that fractions are usually presented without a definition, and described in three incompatible ways all of which are wrong. He then asserts that this is an instance where advanced training is irrelevant to elementary education, because the precise mathematical definition is too complicated. I discuss the general relevance issue in §4. Here I explain that what he describes as the mathematical definition is also wrong, and this is part of the reason it seems too complicated.

Wu says that fractions are equivalence classes of ordered pairs. But this is obviously wrong: the fraction is an ordered pair, not the equivalence class. Describing it as an equivalence class is a sloppy but common imprecision. The first reason for this imprecision is the desire to identify fractions with rational numbers. There are many fractions that represent any given rational. This is not a problem if we think of these as *names* for the number, but if we don't want to distinguish between names and things then we have to think of fractions as equivalence classes. This confusion is more-or-less harmless in a college class when fractions are used to construct the rational numbers. It is not harmless in elementary education where most of the difficulty concerns the *difference* between names and numbers (e.g. determining if two fractions represent the same number).

A deeper reason for the imprecision is that an explicit description of a thing (e.g. a decimal) automatically shows the thing exists and is well-defined. Both of these must be established for implicit definitions. Moreover, problems with things like $\frac{2}{0}$ show that this is not completely straightforward for fractions. Wu's "a fraction is an equivalence class" reflects the usual way to address this: it is short for "things satisfying the defining property of fractions do exist, in the ring of equivalence classes of fractions". This is inappropriate for elementary education in the same way as "$\sqrt{2}$ exists in the quotient polynomial ring $\mathbb{Q}[x]/(x^2-2)$", but neither is necessary. Uniqueness, not existence, is the key issue for fractions, square roots, and certain other implicit definitions. Roughly speaking, the reasons for uniqueness are so robust that they imply existence in an extended context (see *Proof projects for teachers of mathematics* in [7] for more about this). The point is that the feature that Wu sees as making the "real" definition excessively complicated can be omitted without harm, and can be quickly and easily filled in as soon as students are sophisticated enough to notice something is missing.

Another example Wu cites is the irrelevance of modern geometry courses to the Euclidean-style geometry taught in schools. It seems to me that a connection is impossible because Euclidean geometry no longer qualifies as mathematics. It certainly does not meet his "fundamental principles" ([9] p. 378): the objects are not precisely defined; the arguments are not precise and can barely be considered logical; the methods are certainly not coherent with modern mathematics, and it is not particularly goal-oriented. Defects are analyzed in detail in [6] (see the index); I mention only one here. Many Euclidean arguments are essentially proof-by-example. One is supposed to choose "generic" examples that somehow implement the universal quantifier ("for every triangle...") but there is no criterion for when this is successful; it sometimes fails; and it works most of the time only because the subject is so simple. It is a very poor model for mathematical reasoning.

In his discussion of the need for inservice training Wu writes:

> It is time for us to break out of the vicious cycle by exposing teachers to a mathematically principled version of the mathematics taught in K-12.

But there *is* no mathematically principled version of much of what is now taught in K-12: too much is ineffective, obsolete, or plain wrong. Breaking out of the vicious cycle will require profound changes in K-12 teaching. Wu actually gives the argument for this, particularly in his section on "Fundamental Principles", but somehow draws the opposite conclusion.

The above may seem like an attack on Wu's article, but in fact I think he is right more than he is wrong. His "Fundamental Principles of Mathematics", for instance, could be more precise and purposeful but they are sensible and refreshingly free of philosophy.

I would like to expand on some of his objections. First, fractions are *not* the same as ratios. In particular, ratios interact poorly with negative numbers, and if one wants to do vigorous arithmetic then one more-or-less has to give up one or the other (or be very careful and sophisticated). Descarte accepted ratios as the correct division-like operation, and consequently found negative numbers so problematic that he referred to them as "false numbers". Imaginary numbers were less problematic! Ratios are *related* to fractions, but to equate the two is a fairly serious falsehood. The second objection is that the 'parts-of-a-whole' approach is both dysfunctional and specialized to the integers. It might be seen as an *application* of integer fractions, but to present it as the definition seriously misrepresents the concept.

5.2. **Ira J. Papick.** Papick's article [5] is titled *Strengthening the Mathematical Content Knowledge of Middle and Secondary Mathematics Teachers.* On p. 390 he gives a long list of student questions that "teachers should be prepared to address in a mathematically meaningful way." But almost all of these questions reflect serious confusions, misrepresentations, or outright errors in the standard curriculum. Rather than preparing teachers to straighten out students who happen to notice them, shouldn't we straighten out the curriculum so they don't occur? Isn't this better seen as a list of things that need to be fixed in contemporary K-12 mathematics? Some specifics:

- Question 2 reflects confusion due to sloppy and ambiguous use of "="; see §3.1.3.
- Question 3 reflects the confusion between fractions and the rational numbers they represent.
- Question 5 reflects egregious errors in educational use of guessing the next term in a sequence.

- Question 6 on $2^{\sqrt{2}}$ reflects problems that Wu [9] p. 376 addresses with his "Fundamental assumption of school mathematics": formulas for rationals extend (by continuity) to real numbers. This ensures things gotten this way won't be false, but it is not functional as a definition. At some point the miraculous exponential function should be introduced, and the genuinely functional definition $A^B = \mathsf{exp}(A\,\mathsf{log}(A))$ given. Contemplating $2^{\sqrt{2}}$ reveals why this is a good thing.
- Question 7, on the difference between $\frac{(x+3)(x-2)}{x-2}$ and $x+3$, reveals an imprecision that teachers often abuse (and reenforce) in testing. Claiming these are the same, and that the nonsingular version is the "right" form, enables them to mark as wrong an answer that doesn't include the cancellation. Many interpretations of "simplify" are similarly problematic.
- Question 9, on the relevance of the quadratic formula in the age of calculators, reflects deep confusion in the educational community. Calculators belong in the curriculum, but using them in ways that do not undercut long-term learning goals is a much more subtle problem than generally appreciated. Most current programs are counterproductive.

Papick describes a number of courses developed to help teachers deal with such issues. The Algebra for Algebra Teachers mentioned on p. 392, for instance, has a relevant list of topics. But it cannot connect with teaching until educational methodology becomes much more precise and mathematical.

5.3. **Ruth M. Heaton and W. James Lewis.** Their article is *A Mathematician-Mathematics Educator Partnership to Teach Teachers.* Lewis (the mathematician) describes his goal as

> . . . to help teachers become productive mathematical thinkers with a toolbox of skills and knowledge to use to experiment, conjecture, reason, and ultimately solve problems.

They describe his use of the "chicken nugget conundrum", used as an example here in §3.2, in one of these courses. However the description raises questions about their definitions of "mathematical thinker", "toolbox of skills", etc.

First, the students did not use mathematical models. Modeling is an essential part of the professional toolbox, but contemporary educational philosophy rejects it (see §4.2.4). This problem has a subtle mathematical core so massive confusion was a predictable consequence. Another consequence is that the students missed important structure. The key ingredient mathematically is that the coefficients (see§3.2) are nonnegative. In the chicken formulation this is an implicit property of boxes of nuggets, and since the solutions were still phrased in terms of chicken (see Susan's solution, p. 398) this was never made explicit. In effect they learned something about chicken nuggets rather than something about integers. Yet another consequence was that (as the authors observed) the students' explanations were long and wordy. Professionals quit writing things out in word form in the seventeenth century when modeling became widespread.

A second concern is that the students did not use the appropriate mathematical tool. Eventually Susan used divisibility to justify the answer, but this is unwieldy and the extract of her work suggests she could not have found it on her own this way. Why was she not steered toward modular arithmetic during the mentoring sessions? A possibility is that the description of modular arithmetic Lewis sees as appropriate for school use is not actually functional. This is true of Dubinsky; see the next section.

To summarize: the authors present this as an example where success was salvaged from unexpected disaster. But it seems to me the disaster was predictable and the salvage fell considerably short of providing a "toolbox of skills and knowledge". It looks successful only from within mainstream educational theory.

5.4. **Ed Dubinsky and Robert P. Moses.** Dubinsky and Moses present a synthesis of everything relevant, and the Civil Rights movement as well.

Dubinsky writes that at one point he realized that to significantly improve his students' learning he would have to better understand the process of student learning. But instead of studying students, he studied the education literature. When he found Paget he says "...I knew I had come home" [3], p. 402. Some of Paget's insights are impressive, but they are abstract high-level constructs that almost invite abuse at micro levels. For instance

> If one has built appropriate [mental] structures, very early concepts can be grasped easily ... through normal life experiences .... Later, with such structures, more advanced concepts can be learned without undue difficulty via any pedagogical method that relates the concept to the structures. If, however, one does not possess structures appropriate for a concept, it is nearly impossible to learn it.

This implicitly describes two different approaches, and has a very strong hidden implication that they are compatible. First, it logically establishes "learning without undue difficulty" as a criterion for "appropriate structure", and conversely "nearly impossible to learn it" as a criterion for the lack thereof. Logically this should mean that downstream consequences should be a primary concern at all levels, and in particular, elementary education should be highly constrained by the need to support higher levels. I find this entirely reasonable. But the statement also claims "very early concepts can be grasped easily through normal life experiences". The hidden implication is that early concepts grasped this way will support later learning. Educators, including the authors and Beckmann in [2], take this hidden implication as an article of faith. Indeed their faith is so strong that rather than looking downstream they *compartmentalize* levels. Unfortunately, rather superficial examination shows this belief to be false. I illustrate this with Dubinsky's examples.

On page 403 he discusses strategies to help students reconcile the views of $2/3$ as a *process* (parts of a whole) and an *object* that encapsulates the process. But as Wu [9] p. 374, points out, neither of these views is mathematically sound. Perhaps they can be "grasped through normal life experiences" but neither supports later work, e.g. with $(x+3)/(x-4)$. Shouldn't they be discarded and replaced with something functional, rather than reconciled?

On page 407 the authors describe a way to help students grasp modular arithmetic, in the context of division-with-remainder:

$$a = qb + r, \text{ with } 0 \leq r < b$$

They use a game modeled on a clock, with a student walking a units around a circle of length b. The number of cycles is q, and the final position on the circle gives r. The first problem is that this shows "grasp" is interpreted as the "relate to and admire" of philosophy rather than the "exploit as a tool" of mathematics.

In particular the "grasping" this game provides does not provide an "appropriate structure" to support applications such as §3.2, 3.3, and 3.5.4. A more mathematical objection is that modular arithmetic is concerned with equivalence classes obtained by identifying b with 0. This does not mean the equivalence-class concept should be used, but we should be consistent with it and there are two problems here. First, the remainder is a particularly nice representative of the equivalence class, but it invites confusion to identify it with the equivalence class. Second, the quotient q is *not* part of modular arithmetic, and including it makes applications significantly more difficult.

On page 407 Bob describes how "people talk', 'feature talk' and questions about trips on the Red Line in Cambridge relate to mathematics:

> ...students mathematizing these trips acquire powerful metaphors and concepts for addition and subtraction very different from their arithmetic metaphors for these operations ...

Again the objective seems to be the "relate to and admire" of philosophy rather than the "exploit as a tool" of mathematics. Similarly, he cannot have looked downstream: these metaphors interfere with algebra or even fluent work with numbers, so they are more likely to be barriers to be overcome than "mental structures" that support later learning.

Summary: many educators have unquestioning faith that any way they interpret "grasping early concepts through normal life experiences" will automatically support later learning. In most cases this faith is unjustified, and downstream responsibilities are not being met.

5.5. **Mark Saul.** The International Mathematical Olympiad is the subject of Saul's article [8]. It is not clear why it was included in the special issue, because it focuses more on the associated community and some of the mathematicians who have participated than on actual content. However, it provides an opportunity to express concerns about the trick problems used in such competitions.

Trick problems depend on a clever insight or special feature that if missed makes them hard, and if seen makes them easy. But this gives a misleading view of the nature and goals of mathematics, and the activities of mathematicians. First:

- Trick problems are contrived. Small variations typically give identical-looking problems that are impossible with elementary methods; see §3.2.11 for an example. Consequently these are contests between problem designers and students, not between nature and students.
- Many K-12 teachers take the use in competitions to mean that these problems are next step up, and talented students should be challenged with trick problems. Longer and more involved but genuinely illuminating mathematical opportunities go unused.
- Many talented students do not have the necessary quick cleverness, or don't like tricks, and are turned off by this view of mathematics.

Trick problems also make competitions problematic as a recruiting tool for the profession. Saul takes pride in the fact that some outstanding mathematicians were first identified through their performance in competitions. But most high-scoring competitors did not become mathematicians. In fact most mathematics, deep work especially, is slow and methodical rather than quick and clever. Further, the everyday power of mathematics is that persistence, appropriate techniques, and good work habits succeed with long, hard problems where quick cleverness is powerless. Quick and clever people often find the long-haul tenacity required for real accomplishment either unattractive or impossible.

The other side of this coin is that most outstanding mathematicians did not distinguish themselves in competitions. Again, the professionally productive methodical and persistent mindset does not correlate well with the quick cleverness needed for competitions. I was put on my college Putnam team because my teachers thought I had the greatest mathematical potential in the class. They were right about professional accomplishment but wrong about the Putnam: I had the lowest score on the team by far, and may have kept it out of national rankings.

Summary: no doubt trick problems have their place, and competitions may require them. However they should not be presented as representing real mathematics, should not be the default next step for promising high-school students, and should not be represented as ideal recruiting tools for either mathematicians or potential users of mathematics.

References

[1] Special Issue on Education, *Notices of the American Mathematical Society* **58** (March 2011).

[2] , Beckmann, Sylvia; *The Community of Math Teachers, from Elementary School to Graduate School*, pp. 368–371 in [1].

[3] Dubinsky, Ed; Moses, Robert P; *Philosophy, Math Research, Math Ed Research, K-12 Education, and the Civil Rights Movement: A Synthesis*, pp. 401–409 in [1].

[4] Heaton, Ruth M; Lewis, W. James; *A mathematician-mathematics educator partnership to teach teachers*, pp. 394–400 in [1].

[5] Papick, Ira; *Strengthening the mathematical content knowledge of middle and secondary mathematics teachers*, pp. 389–393 in [1].

[6] Quinn, Frank; *Contributions to a science of contemporary mathematics*, preprint; current draft at http://www.math.vt.edu/people/quinn/.

[7] ______ *Contributions to a science of mathematical learning*, preprint; current draft at http://www.math.vt.edu/people/quinn/.

[8] Saul, Mark; *More than a System: What We Can Learn from the International Mathematical Olympiad*, pp. 410–416 in [1].

[9] Wu, H., *The mis-education of mathematics teachers*, pp. 372–384 in [1].