The Uses of Spurious Proofs in Teaching Mathematics

The Uses of Spurious Proofs in Teaching Mathematics

Ethan A. Fenn

A Thesis in the Field of Mathematics for Teaching

for the Degree of Master of Liberal Arts in Extension Studies

Harvard University

May, 2018

Abstract

A spurious proof is a mathematical proof that seems to be logically cogent at every step, but reaches a conclusion that is clearly impossible. This occurs because a step of the proof has been cleverly written to conceal its own falsehood, and the hidden falsehood ultimately creates the incongruous conclusion. In the past, close reading of spurious proofs in order to discern which step is the false one has been a niche endeavor of recreational mathematics, and only occasionally used in the classroom as one among many types of long-form problems.

However, spurious proofs have several distinctive—though sometimes neglected—values in the educational context. One of these sources of value is spurious proofs' creation of "cognitive conflict" that stimulates critical thought in the contexts of both general problem solving and learning the proof process. Another underappreciated source of value is spurious proofs' potential fitness for developing the "productive disposition" aspect of mathematical proficiency that sees math as a sensible and useful endeavor. This aspect of proficiency is widely recognized in the literature, but just as widely overlooked in the classroom, likely because it is usually something that must be taken or rejected wholesale. Spurious proofs, though, give teachers a rare means to reinforce this value in their students through an active problem-solving process. With a spurious proof, the coherency of the mathematical system is challenged when the proof offers a result that

contradicts well-known mathematical knowledge, but still seems to be logically supported; the student must then actively validate the system's coherency by searching out the proof's logical flaw. In this way, a student is actively engaged in affirming the consistency of the mathematical system rather than simply accepting it, or not.

After presenting some background information on the history and typology of spurious proofs, this thesis explains their value from cognitive conflict and coherency reinforcement in more detail, and then offers examples of how they may be put to use in the classroom beyond their generic role of simple long-form problems.

Ride With Youth

Speed is the watchword of youth today,
    Speed is the goal and life the stakes;
Down the road and the devil to pay,
    Watching the youngster the oldster quakes;
    Foot on the throttle and off the brakes,
Heedless of chance of skid or spill—
    This is the comment Edison makes:
"Ride with youth if you want a thrill!"

Harsh and bitter the words we say
    Sneering of shebas and sheiks and snakes;
What is it coming to, anyway,
    This generation of queens and rakes,
    Hip-flask toters and free-love fakes?
Nothing of good and all of ill!
    Let's quit croaking, for all our sakes—
Ride with youth if you want a thrill.

Yes, we grumble at youth at play,
    The risk it runs and the chance it takes,
Letting the future bring what it may,
    Making of fortune ducks and drakes.
    Grouchily nursing rheumatic aches
We keep on damning the young—and still
    Deep in our conscience envy wakes:
Ride with youth if you want a thrill.

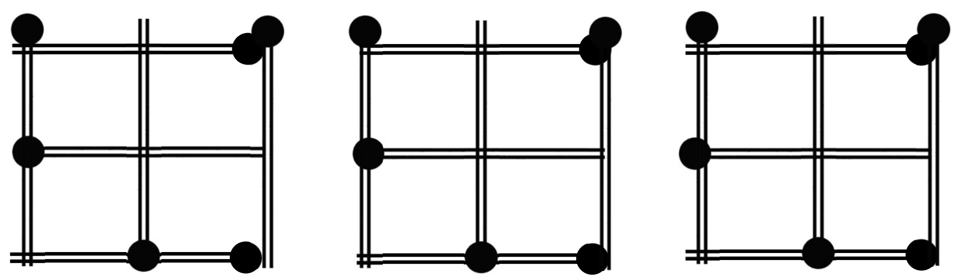
. . .

Prince, desist from your shrugs and shakes;
    Give 'em their fling and pay the bill!
Edison says—and he knows his cakes—
    "Ride with youth if you want a thrill."

F. Gregory Hartswick (1929)

To my father, who got me started in mathematical shenanigans with this:

Take away eight matches and leave eight.



Acknowledgments

The author wishes to thank Andy Engelward and Oliver Knill for many helpful comments along the way. Thanks are also due to the Harvard library staff, especially Rebecca Martin and Anna Rakityanskaya for assistance with Russian sources.

Table of Contents

List of Tables

List of Figures

Chapter I.

Introduction

Spurious proofs—mathematical arguments that seem to be logical at every step of the way, but reach conclusions that are clearly impossible—have long been a niche endeavor of recreational mathematics and occasionally the classroom, as curious amateurs or select groups of students would try to sort through the argument to find the one concealed false step that allows the incongruous conclusion to be reached. When used in the classroom, spurious proofs have merely served as one type of long-form problem among the many available to teachers looking to stretch their students' creative and critical mathematical thinking skills.

In fact, though, spurious proofs have a certain educational values all their own. For one, they are a particularly good generator of cognitive conflict, a temporary state of confusion that can stimulate students to think critically about what has caused their confusion and then restructure their mental framework on a more stable footing; this phenomenon is a useful tool both in general long-form problem solving and in teaching the proof process. In addition, spurious proofs are one of the few sorts of problems that can actively reinforce a key strand of mathematical proficiency: the productive disposition by which students see mathematics as a sensible and cogent endeavor, and see themselves as capable of engaging with it. These two values are fleshed out as follows in the respective sections below:

Chapter II provides some context for the educational theory that follows by reviewing some background information on spurious proofs. This background information includes a formal and precise definition of spurious proofs, along with a close look at how they differ from some similar types of mathematical problems and arguments that involve an element of error, trickery, or confusion. It also covers some history of spurious proofs, and a general typology.

Chapter III will then address the first broad use of spurious proofs in teaching math: the generation of cognitive conflict.

After that, Chapter IV looks at the other potential use of spurious proofs in the math classroom: building a productive disposition by reinforcing the perception of math as a cogent and comprehensible system. In theory, spurious proofs would seem to have a unique potential role in reinforcing the "productive-disposition" aspect of mathematical proficiency by which students see mathematics as a cogent and worthwhile enterprise, and themselves as capable practitioners of it. This aspect of proficiency was one of five interlocking strands introduced in the National Research Council's (2001) report *Adding It Up: Helping Children Learn Mathematics*. That five-strand framework has been extremely influential in the mathematical-education literature since then, but the productive-disposition strand has been largely left out of the party. It is the only one of the five strands related to mathematical attitudes rather than mathematical performance, with the result that research on how to develop a productive disposition in students has lagged behind similar research for other areas of mathematical proficiency. What research there is, however, suggests that mathematical content is an important component of fostering a productive disposition; it is not merely a matter of bolstering self-esteem through the right psychological interaction between student and teacher. Furthermore, spurious proofs fit many of the particular criteria that have been studied for particular exercises to build productive disposition. Finally, all of this ignores the seemingly obvious point that spurious-proof problems are uniquely constituted in that their entire thrust is a challenge to the consistency of the mathematical system, which challenge students must then resolve through an active thought process. Spurious proofs would thus seem to be extremely distinctive in being calibrated directly toward major aspects of productive disposition. For all this potential advantage, though, the use of spurious proofs to boost productive disposition has been under-examined to date.

Chapter V offers a couple of short steps toward a remedy for this, in the form of two spurious-proof exercises for the classroom. Seeing that most of the foregoing discussion centers on the value of standard spurious-proof problems in which the teacher presents the proof to the students in order for them to find the false step, the first exercise fits this pattern, although the particular tangram paradox utilized appears to be original to the present author. The second exercise is a bit more creative, calling on the teacher to guide students through the composition of a spurious proof based on Zeno's paradox of the Dichotomy.

Chapter II.

Background on Spurious Proofs

Before delving in to the educational value of spurious proofs, it will be useful to set out some general background material on their definition, history and typology. Along the way, each of the parts of this chapter will afford the opportunity to compare spurious proofs with various other types of mathematical chicanery with which they share certain similarities, and with which they are often lumped together in collections of recreational problems. These similar mathematical oddities include crankery, errors, ambiguities, paradoxes, riddles, and howlers.

Definition

As noted above, “A spurious proof is a mathematical proof that seems to be logically cogent at every step, but reaches a conclusion that is clearly impossible.” A full definition might expand a bit, so as to read “A spurious proof is a mathematical proof that seems to be logically cogent at every step but through the concealment of one or more false steps reaches a conclusion that is clearly impossible, and is deliberately presented by its author as a problem for resolution by the reader.”

This expanded definition contains two additions: “through the concealment of one or more false steps” and “is deliberately presented by its author as a problem for resolution by the reader.” Together, these additions serve two purposes. First, they describe how and why spurious proofs typically appear in the literature. In a genuine proof, each step depends logically upon the last; the point of presenting a spurious-proof problem is to illustrate that a falsehood in the middle then creates a falsehood at the end. Indeed, a small falsehood can sometimes imply a very large one, once logical consistency has been entirely discarded. Since one false step is enough to accomplish the task of a spurious proof, the possibility of multiple false steps might seem to be overkill in a well-constructed problem; but, as will be seen below, there is at least one notable class of spurious proofs that possess this property.

The second purpose of these additional qualifications in our definition is to distinguish a deliberately constructed spurious proof from several other types of "fake math" (this term is due to Oliver Knill in conversations with the author) including crank efforts, errors, ambiguities, paradoxes, riddles and howlers. All these various types of fake math involve some putative mathematical result that is at variance with established mathematical knowledge, but they vary in the author's awareness of the result's incorrectness, and in how the result is presented to the reader.

Because of this similarity, instances of the various sorts of fake math are often packaged together in recreational collections, in which spurious proofs are sometimes also identified as 'sophisms' or 'fallacies.' It is also a common terminological tack to lump spurious proofs into the general category of paradoxes. Precisely how these various mathematical tricks differ from each other is explained below.

Comparison to Crankery

A crank argument is one that honestly purports to have overturned extremely well established mathematical wisdom—such as claimed solutions to squaring the circle or finding a rational expression for π. Like a spurious proof, a crank argument is a facial proof of something that is obviously false; but a crank effort differs from a spurious proof in that the crank author believes his argument to be true, while a spurious proof's author is aware of the proof's ultimate falsehood and consciously presents it in order to amuse or educate the audience. A relatively intricate crank argument might be turned into a spurious proof by a subsequent reader: If a reader (unlike the crank author) realizes that the conclusion is false, but sees that picking out the particular flaw in the author's argument is a bit of a challenge, then the reader might repackage the crank argument as a spurious proof.

Comparison to Honest Errors

Another key feature of a crank argument is that its author ought to know that it is false, given that the conclusion runs counter to solidly established mathematics. This distinguishes a crank argument from an honest error by a trained and diligent mathematician. In an erroneous mathematical argument, the conclusion is false but believed true by the author, just as with a crank argument. However, with a genuine error, the author is operating on the edge of mathematical knowledge so that the error could understandably slip by both the author and (for a time) the larger mathematical community, even with the appropriate training and diligence being brought to bear on the matter. In some cases, it might be a close call whether a particular mathematical miscue constitutes crankery or an honest error, but a rather good definitive characteristic of cranks was offered by De Morgan (1915, p. 4), which is surely the most exhaustive study of them: "The manner in which a paradoxer will show himself, as to sense or nonsense, will not depend upon what he maintains, but upon whether he has or has not made a sufficient knowledge of what has been done by others, *especially as to the mode of doing it*, a preliminary to inventing knowledge for himself." (Emphasis is original.) In context a 'nonsense paradox' would in De Morgan's (1915, p. 2) terms be a "crotchet" and in modern terms be 'crankery.' Thus, De Morgan would hold that the distinction between crankery and an honest error is that the crank has failed to conduct due research into the existing state of knowledge on his subject matter before offering his argument.

Some famous examples of errors in mathematical history have included early putative proofs of the Four-Color Theorem and Fermat's Last Theorem (Bell, 1937, pp. 472–73; Thomas, 1998, p. 848), both of which have only been definitively proven in recent times and with more intricate methods. Another error has garnered some attention in recent years because of the political ambitions of its author: Daniel Biss was once on the mathematics faculty at the University of Chicago. At around the same time that he began a run for the Illinois House of Representatives, it was discovered that his doctoral thesis and several publications that grew out of it relied on a subtle but fatal error (Anderson, 2007; Dee, 2007; Henderson, 2007; Mnev, 2007; Szpiro, 2010). Certain parties considered this an issue in Biss's initial run for the state legislature (Dee, 2007), arguing in effect that the error betrayed a large enough lack of diligence to cross the line from honest to crankery. Those with direct knowledge of the situation disputed that assessment, however (Anderson, 2007). Biss lost his initial run for office (Szpiro, 2010, p. 98), but he was eventually elected to the state house and then to the state senate (Sfondeles, 2017). Nonetheless, with Biss now seeking the Democratic nomination for governor, the old issue of his mathematical error has arisen again (Sfondeles, 2017).

Like a crank argument, a mathematical error can be turned in to a spurious proof by a subsequent reader who, once the error has been discovered, presents the argument to others to see if they can find it as well. Indeed, since mathematical errors have, by definition, fooled some number of professional mathematicians for a length of time, they can be fruitful ground for the sort of subtle and easily concealed error that makes for a good spurious-proof problem.

Comparison to Ambiguity

A mathematical ambiguity is a problem that has at least two different plausible interpretations leading to different results. Very often, only one interpretation is conceived originally by the author, and the ambiguity is discovered only when one or more readers lands on the other interpretation. In a typical example, a teacher or examining board puts an ambiguous question on an exam, and there is a great kerfuffle when a mass of students give a reasonable answer that is originally marked as wrong. A large class of ambiguities center on varying interpretations of the order of operations, as slightly different orders have historically been embodied in different calculators, spreadsheets, explicit rules and implicit conventions. Generally speaking, a clear problem with multiple solutions, like $x^2 = 9$ with solutions $x = \pm 3,$ does not constitute an ambiguity; the multiple solutions must arise from disjoint interpretations of the problem. However, in the context of a younger student who is accustomed to problems having exactly one answer, a multiple-solution problem might fairly be dubbed an ambiguity because (in that student's learning context) the problem is not, in fact, clear.

Comparison to Paradoxes

A third type of mathematical misadventure that must be distinguished from a spurious proof is the paradox. In a paradox, the author presents the reader with a genuine argument that contravenes established mathematics, but with some uncertainty for both the author and reader as to whether the author has erred or whether established mathematics has been definitively overturned. The author might present the paradox with the hope of persuading the reader to the author's point of view, or with the hope that a reader might find a resolution that has so far eluded the author, or simply in order to prompt further discussion of the subject matter.

One famous paradox still generating discussion in the philosophy of mathematics is Russell's Paradox: If a set is defined to contain precisely those sets that do not contain themselves, does it contain itself or not? Put more operatically, if the Barber of Seville shaves only those who do not shave themselves, who shaves the barber? If the set does contain itself, then it is outside of its definition, which means it can't contain itself; but if it doesn't contain itself, then it is within its definition, and must contain itself.

Some additional famous paradoxes are those put forth by the Greek philosopher Zeno of Elea. Zeno's paradoxes are described in more detail below at 13, but his basic 'paradox paradigm' was to argue that some example of motion or progress is impossible because it must comprise an infinite number of intermediate stages, each very small but greater than zero (Heath, 1921, pp. 274–81). Zeno and all his readers throughout history have known that for motion to be impossible flouts common sense, but whether Zeno or common sense was correct was not obvious to the Greeks of his day.

At the level of a philosophical conundrum, Zeno still appears to inspire some lively discussion today, although if one interprets his arguments mathematically, it is now clear where they break down (see below at pp. 13–17). This leads in to a commonality between paradoxes and other forms of fake math that we have examined: they can be turned in to spurious proofs by a subsequent reader who sees where the miscue is, and presents the paradox to further readers to find the miscue as an exercise. Since a paradox, by definition, has caused some confusion amongst experts in the past, it would seem to be fruitful ground for an easily camouflaged false step that would undergird a spurious-proof problem.

Comparison to Riddles

Another category of fake math that bears some comparison to a spurious proof is the riddle. A riddle is similar to a spurious proof in that the author presents the reader with a mathematically incongruous conclusion, with the author aware of the incongruity and intending for the reader to resolve it as an exercise; the difference between a riddle and a spurious proof is that with a riddle, the author asks the reader to devise the entire incongruous argument on the reader's own, while with a spurious proof, the author intends for the reader to find the single camouflaged false step in an argument that the author supplies. To illustrate the difference, consider the famous riddle 'Prove that Halloween equals Christmas.' The solution can be rendered in one string equality as

$$\begin{aligned} \mathit{Halloween} &\equiv \text{Oct. } 31 = {}_{8}31 = 3 \cdot 8 + 1 \cdot 1 = 2 \cdot 10 + 5 \cdot 1 = {}_{10}25 \\ &= \text{Dec. } 25 \equiv \mathit{Christmas}. \end{aligned} \tag{II.1}$$

The trick that allows the absurd conclusion to be reached is the double meanings of Oct. 31 as both 'October 31' and 'Octal 31,' and Dec. 25 as both 'December 25' and 'Decimal 25.' This trick is quite obvious once the argument is spelled out, so the problem wouldn't work as a spurious proof; instead, the writer asks the reader to find the trick argument from scratch, which makes the problem a riddle.

Sometimes, one can make a riddle out of a paradox: What Bertrand Russell say when he discovered his famous paradox? 'I'm so happy I can't contain myself!'

Comparison to Howlers

A last type of mathematical trickery that one sometimes encounters in the recreational-math literature is the howler. A howler is a mathematical argument that reaches a correct result by way of an absurdly errant process (Maxwell, 1959, pp. 9–10). A howler is usually not suitable as a problem for reader resolution: the answer is correct; and, for the error to be interesting, it typically must be egregious enough to be obvious. However, howlers do often make for amusing reading and, since they rely on a mishap, they are often packaged with spurious proofs in collections.

Comparison to Problems

Finally, it should also be noted, if one is reading through a recreational collection and trying to classify its material, that all these types of fake math are to be distinguished from a straightforward problem. In this last case, there is no error or contradiction—real or apparent—in the math; it merely takes the reader some creative thought to reach a solution.

That said, however, an old problem (Verschaffel, Greer, & De Corte, 2000, pp. 3–6) that has recently surfaced in the news media (BBC News, 2018; Rezaian, 2018) does suggest a distinct form of mathematical trickery: the unsolvable problem, in which the problem's author gives the reader a problem that is not capable of being solved from the information provided, with the purpose being to stimulate thought on the reader's part. Verschaffel, Greer & De Corte (2000) report that the format elicits a remarkable tendency for students to attempt a precise solution when it is clear that none is possible, a fact that they attribute to students analyzing exercises through the lens of expected classroom conventions (which an unsolvable problem transgresses) to the exclusion of ordinary common sense.

History

To a large extent, the history of spurious proofs consists of their being shared through informal circles of recreational mathematics for many years, and, indeed throughout the history of mathematics. Given the extent of this type of sharing, an exhaustive review of every instance of the mention of such proofs would be beyond the scope of this project. However, there are a few high points that merit mention:

Zeno

The first, and probably still most famous spurious proofs, were produced by the ancient Greek philosopher Zeno of Elea in the fifth century B.C. Zeno presented a number of philosophical conundra (the four chief ones being known as Achilles and the Tortoise, the Arrow, the Dichotomy, and the Stadium); they all in various ways centered around the proposition that discrete motion or progress toward a goal is impossible because it must involve infinitesimally small intermediate stages, and it is surely impossible for a discretely discernible quantity to be made up from pieces that are all equivalent to zeroes. Although originally presented as philosophical paradoxes, Zeno's arguments can also be seen as spurious proofs of various mathematical topics related to limits and infinite series.

A good general overview of Zeno's work can be found in (Heath, 1921, pp. 271–84). However, given the importance of Zeno in the history of spurious proofs, and the continuing popularity of his paradigm in contemporary spurious-proof problems, it is worth delving into a more specific illustration of Zeno's method, and a good example is his paradox of Achilles and the tortoise (Heath, 1921, pp. 275–76).

The paradox of Achilles and the Tortoise proposes that if the mighty hero Achilles were to run a race against a plodding tortoise, with the tortoise given a head start to compensate for Achilles' superior speed, the hero would never catch up with his reptilian competitor because he would have to traverse an infinite number of intermediate intervals in the interim. While Zeno offered his argument as a pure philosophical conundrum, it is presented here in mathematical terms in order to illustrate how it operates as a spurious proof.

Thus, we begin with Achilles racing a tortoise, and with the tortoise being given a head start. Mathematically, let us say that this head start, or gap, is a distance of g_0. Let us also assume that Achilles runs with a constant speed s_A, and that the tortoise 'runs' with a constant speed s_T.

If Achilles is ever to catch up with the tortoise, he must first traverse the distance of the tortoise's initial head start. Numerically, we can calculate the time it takes him: since distance equals speed multiplied by time, for Achilles to traverse the head-start distance g_0 will require a time t_1 equal to the head-start distance g_0 divided by Achilles' speed s_A; that is, $t_1 = g_0 / s_A$.

However, in the time that Achilles takes to make up the initial head start, the tortoise will have advanced a bit further to maintain his lead. We can calculate this new gap g_1 by noting that it is the distance covered by the tortoise moving at speed s_T for time t_1. Again using the familiar formula distance equals speed multiplied by time, we have $g_1 = s_T \cdot t_1$. Furthermore, by substituting for t_1, we can calculate the relationship between the new gap and the old gap:

$$g_1 = s_T \cdot t_1 = s_T \cdot \frac{g_0}{s_A} = \frac{s_T}{s_A} \cdot g_0. \tag{II.2}$$

However, in the time it takes Achilles to make up this new gap, the tortoise will have advanced still more. Indeed, for every gap Achilles closes, the tortoise opens up a new one. Mathematically, we can use the same logic behind Equation

(II.2) to infer that $g_2 = (s_T / s_A) \cdot g_1$ and, in general, $g_n = (s_T / s_A) \cdot g_{n-1}$. Iterating *n* times on this last equation,

$$g_n = (s_T / s_A) \cdot g_{n-1} = (s_T / s_A)^2 \cdot g_{n-2} = \ \ldots \ = (s_T / s_A)^n \cdot g_0. \tag{II.3}$$

Since Achilles is faster than the tortoise, we know that the factor s_T / s_A is less than one, so the tortoise's head starts are decreasing, but always greater than zero.

From here, Zeno would have us believe that Achilles will never catch up with the tortoise since, for every gap Achilles closes, the tortoise opens up a new one. However, putting the argument in mathematical terms lets us fairly quickly see one error in Zeno's reasoning that would not have been apparent in the days of ancient Greece.

Modern mathematics tells us that an infinite series of terms may nonetheless have a finite sum if the terms converge to zero sufficiently quickly, and this is precisely what happens here: Zeno infers that the tortoise runs forever without being caught, but the distance *D* that the tortoise runs before being caught must equal the sum of all his head starts, and we can set out a formula for this:

$$D = g_0 + g_1 + g_2 + \ \ldots \ = g_0 + g_0 \cdot \frac{s_T}{s_A} + g_0 \cdot \left(\frac{s_T}{s_A}\right)^2 + \ \ldots \ = g_0 \cdot \sum_{n=0}^{\infty} \left(\frac{s_T}{s_A}\right)^n. \tag{II.4}$$

A modern mathematician sees that this is a geometric series with a term ratio less than one, which is known to converge to a finite sum of the initial term divided by one minus the term ratio, so that

$$D = \frac{g_0}{1 - s_T / s_A}. \tag{II.5}$$

Thus, one false step in Zeno's argument comes from the premise that the sum of an infinite number of terms must be infinite—a premise that would not be disproven until later in the history of mathematics.

However, the continued philosophical interest in Zeno arises from a second fallacy that arises in his argument. In mathematical terms, we might call this fallacy the Limit Fallacy, and it is defined by extrapolating some property from the individual terms of some infinite sequence to the limit of that sequence, or vice versa. In the case of Zeno, the Achilles and the Tortoise paradox illustrates the Limit Fallacy as follows: Giving the tortoise a fixed head start, we can in fact figure out at what point Achilles will catch the tortoise, using the formula of Equation (II.5). Given the speeds of Achilles and the tortoise, we can then determine what (finite) time will elapse before that moment. That interval of time can be partitioned into subintervals; and at any partition point, we can determine where Achilles and the tortoise are on the course based on their speeds and starting positions. However, any such actual partition can only have a finite number of partition points. We can imagine a sequence of partitionings, each with one more partition point than the last, that grow to look more and more like an infinite geometric sequence as we extend them further; but ultimately, that infinite sequence is a limit, a mathematical construct that mathematicians can describe and manipulate in certain ways if we know how, but not something that can directly characterize a phenomenon in the natural world. Zeno looks to the idealized end state of an ever finer sequence of partitionings of time and space, sees that some spans of time or space there must be infinitesimal, and declares a paradox because an actual span of time or space cannot be infinitesimal. However, he is conflating the infinitesimal properties of that idealized end-state partition with the properties of the actual real-life partitions that mark the way there, and those do not present any problem of infinity.

Zeno's perpetration of two distinct false steps make his paradoxes into somewhat odd cases when they are utilized as spurious proofs. However, this oddity is only empirical, not logical. If a core premise of spurious proofs is that one false step is sufficient to produce an absurd result, then two false steps will also do nicely.

19th Century Recreational Math: Loyd and Dudeney

Spurious proofs received a boost in interest from an upsurge in recreational math starting in the late 19th century and extending in to the 20th. The leaders of this upsurge were Sam Loyd (S. Loyd, 1903, 1914, 1959, 1960) in the United States and Henry Ernest Dudeney (1908, 1917, 1967) in Great Britain. In their various works, each presented a wide variety of mathematical puzzles, including amusingly framed problems and various sorts of mathematical trickery, particularly riddles and a few spurious proofs. For the latter, see for instance Gardner (1956, Chapter 8). Within the field of spurious proofs, they each made a bit of a specialty out of seemingly incongruous rearrangements of tangram tiles (discussed in more detail below at p. 27), capitalizing on a cultural trend that had been popular since at least (Goodrich, 1817).

Loyd and Dudeney also exemplified to an extreme degree the informal style of citation and acknowledgment that tends to characterize recreational as compared to academic mathematics. Martin Gardner, the modern puzzle guru who has edited works by both Loyd and Dudeney, concludes that

> In some cases, it is possible to say with certainty that Dudeney borrowed from Loyd; in other cases, that Loyd borrowed from Dudeney. The task of tracing individual puzzles to their first publication by either man is so formidable, however, that one hesitates to say which expert took the most from the other. There was considerable rivalry between the two puzzlists while they were active (only once in the entire *Cyclopedia* [of Loyd] is Dudeney's name even mentioned), and apparently each did not hesitate to appropriate and modify the other's inventions. In addition, both men drew heavily on common sources—traditional puzzles to which they gave new twists, and new puzzles of anonymous origin that passed from person to person in the manner of new jokes and limericks (S. Loyd, 1959, p. xv) (editor's introduction).

Loyd, in particular, has become notorious for later claiming credit for puzzles with which he had no documented connection at the time of their invention (Slocum, 2006; Slocum & Sonneveld, 2006, pp. 75–109).

De Morgan

Another early compendium of various mathematical fallacies was Augustus De Morgan's *A Budget of Paradoxes* (De Morgan, 1915). De Morgan's volume mainly collected the work of genuine cranks who thought they had really solved impossible problems like squaring the circle or finding a rational expression for π, but it also notes a few spurious proofs in the conscious sense as treated in the present paper.

Russia

There seems to have been a surge of interest in spurious proofs in Russia during the early part of the 20^{th} century. Various Russian sources from the period are reported in Bradis, Minkovskii, and Kharcheva (1963), which seems to be one of few such sources translated and published in the West. Another is Dubnov (1963).

Lyamin (1911) is available in the West in its original Russian, and, together with other work by the same author, it may shed some light on a source of appeal for spurious proofs. Though the combination may seem surprising, at least to the modern American reader, Lyamin (1903) also wrote on various topics in the area of mysticism. For Lyamin and his audience—and perhaps more for modern readers than we realize—part of the appeal of spurious proofs seems to have been that, just for a moment, they offer a suggestion that the world around us has more to it than it seems.

Modern Recreational Math

Though spurious proofs were not prominent for some years after the very early 20th century, they enjoyed a revival as part of a general resurgence in recreational math that began in the 1950s and '60s, and continues to the present day. Though a full explanation of why recreational math fell into and out of the public consciousness is well beyond the scope of this paper, one might make a starting conjecture that two world wars and an economic depression left people with more important things on their mind; but that after World War II, a general economic upsurge, a rise in college education fueled by the economy and the G.I. Bill, and the attention to math and science prodded by the cold war all prodded the public to take more of an interest in mathematics for fun.

Beginning in the '50s and '60s, old works of recreational math were reprinted (Bradis et al., 1963; Dubnov, 1963; Dudeney, 1908, 1917, 1967; Fetisov, 1963; S. Loyd, 1903, 1914, 1959, 1960; Lyamin, 1911; Northrop, 1944) and new ones were written (Edward J Barbeau, 2000, 2013; Bunch, 1997; Farlow, 2014; Fixx, 1972; Gardner, 1988a, 1988b, 1956, 1978, 1982, 1983, 1987, Huck & Sandler, 1984a, 1984b; Klymchuk & Staples, 2013; Maxwell, 1959; McInroy, 1994; Movshovitz-Hadar & Webb, 1998; Nufer, 2017; Read, 1965; Smullyan, 1978; Van Note, 1973). The foregoing lists are far from complete, at least relative to the general universe of recreational math; but they give a pretty full view of the sources with a focus on spurious proofs.

Types

In principle, there are two obvious classification schemes by which one might categorize spurious proofs: by the branch of mathematics involved, or by the type of fallacy perpetrated. To group spurious proofs by branch of mathematics is a common practice in published collections, but it would not add much to the analysis here. At least two prior efforts have been made at classifications by fallacy type, but each has its drawbacks, and a system may be imagined that is fuller and more consistent in its levels of classification.

Kondratieva (2009b, pp. 65–67) lumps together as 'paradoxes' what this analysis defines separately as "paradoxes" and "spurious proofs," and posits a very broad classification scheme of self-referential "Semantic paradoxes[,]" "Paradoxes involving limiting processes and infinite sets[,]" "Paradoxes resulting from flawed reasoning, arithmetic error, or faulty logic" (with which Kondratieva (2009b, p. 68) seems to lump all geometric errors), and "True statements which contradict common intuition[.]" This classification system is displayed in the center section of Table II.1 below.

In addition, Bradis et al. (1963, pp. 6–37) posit a classification system of spurious proofs, strictly defined, in accord with the present analytical framework. Their system is set out below in the left-hand section of Table II.1 (with some paraphrasing), and it is a good jumping-off point for laying out a full system of classification. However, when one considers all of the examples of spurious proofs that one encounters, there are several common categories that they omit, and some they list separately that would do with some consolidation.

Taking into account both Bradis et al.'s (1963) framework and examples that seem to fall outside it, one can set out a classification system as follows. These categories are also reflected above in the right-hand section of Table II.1.

Wordplay

Bradis et al. (1963) list four items among their top-level categories of fallacy that might all be classified together under the term 'Wordplay.' Specifically, they list "Incorrectness of speech" with five subcategories (1963, pp. 9–12); "Errors which are the consequence of a literal interpretation of the abbreviated conventional formulation of some geometric propositions" (1963, pp. 31–33); "Violation of the sense of conventional notations" (1963, pp. 33–34); and "Deviation from the thesis" (1963, pp. 34–37).

The five subcategories under the first heading all involve some form of verbal or symbolic ambiguity. Of particular note is the symbolic ambiguity, which they term "Ambiguity of construction" (1963, p. 11), since it includes the frequently encountered phenomenon of confusion over the order of operations. The

Table II.1

Classification of Spurious Proofs

<table>
<tr><th colspan="2">(Bradis et al., 1963)</th><th>(Kondratieva, 2009b)</th><th>Suggested</th></tr>
<tr><td colspan="2"></td><td>Limits & Infinity</td><td>limit fallacy</td></tr>
<tr><td colspan="2">Extension to Exceptional Cases</td><td rowspan="10">Errors</td><td rowspan="4">Misapplication</td></tr>
<tr><td colspan="2">Extension to General Case</td></tr>
<tr><td colspan="2">Assumption of Converse</td></tr>
<tr><td colspan="2">Replacement of Definition by Intuition</td></tr>
<tr><td rowspan="6">Misdrawn Diagram</td><td>coincident points assumed distinct;</td><td rowspan="5">Misdrawn Diagram</td></tr>
<tr><td>distinct points assumed coincident;</td></tr>
<tr><td>impossible location;</td></tr>
<tr><td>impossible intersection;</td></tr>
<tr><td>broken line assumed straight;</td></tr>
<tr><td>straight line assumed broken</td><td>Rearrangement</td></tr>
<tr><td>Incorrectness of Speech</td><td>word definition;
pronunciation;
construction;
distribution;
composition</td><td rowspan="4"></td><td rowspan="4">Wordplay</td></tr>
<tr><td colspan="2">Imprecise Colloquialism</td></tr>
<tr><td colspan="2">Misuse of Notation</td></tr>
<tr><td colspan="2">Deviation from Thesis</td></tr>
<tr><td colspan="2"></td><td>Semantic Paradoxes</td><td></td></tr>
<tr><td colspan="2"></td><td>Counterintuitive Truth Paradoxes</td><td></td></tr>
</table>

Note. First-level categories are roman, capitalized, and centered. Second-level categories are roman, capitalized, and right-aligned. Third-level categories are italic, lower-case, and right-aligned.

second top-level heading could presumably be extended to include other instances where there is ambiguity introduced by a colloquial formulation of a theorem, not just instances from the geometric context. The third is meant to capture examples such as cases where a generic series, say, is illustrated by five terms and an ellipsis; but particular parameter values imply that the series will have fewer terms—say, three. If these parameter values are plugged in to the five-term expression of the series without thinking, an ambiguity results. Finally, "Deviation from the thesis" refers to examples where a premise is proven that does not quite mean the same thing as the conclusion proffered at the outset, or where the proof proceeds from a premise that does not quite mean the same thing as the original hypothesis.

All these cases involve some sort of ambiguity of words or symbols, so it makes sense to lump them all into a single category, which is designated here as 'Wordplay.' Many forms of ambiguity are easy to spot once set out in a formal proof, so they tend to work better in other forms of mathematical trickery, like riddles, rather than spurious proofs. "Ambiguities of pronunciation"—that is, puns—typically fall into this category. Other forms of ambiguity, though, such as with the notation in a series-with-ellipsis case, can often be camouflaged readily enough to serve as the basis for a spurious proof.

Misapplication

There are four more top-level categories in Bradis et al. (1963) that share the common essential feature of applying some mathematical principle where it does not really belong: "Extension to exceptional cases (1963, pp. 12–13); "Ascribing properties of a particular form to the whole species" (1963, pp. 13–14); "Incorrect applications of the principle of direct deductions by the converse" (1963, pp. 14–18); and "Replacement of precise definitions by geometric intuition" (1963, pp. 18–21). The first category refers to proofs that apply a principle in a case that is actually an exception to the principle, such as the common algebraic proofs where there is a hidden division by zero. The second category is essentially a form of the first: it also involves utilization of a principle in cases where the principle does not hold, only those cases are now a broad category instead of a trivial exception. The third category is pretty self-explanatory. Finally, the fourth category is meant by the authors to refer to cases where a diagram seems to display some geometric property, but the property cannot really be proven and is not true. These cases would seem to be common enough, but one could also imagine proofs based in other areas of math where some class of objects appeared all to have some property, but in fact did not; potentially, then, the category might be made more general. Since all these categories involve instances of some mathematical principle being applied in a case where it really does not hold, it makes sense to consolidate them together under the heading of 'Misapplication.'

It is worth noting that there are other potential categories of misapplied principle which are not enumerated by Bradis et al. (1963) but still would fall in the general category of misapplication. Indeed, we have already seen one.

Limit fallacy. Specifically, recall the Limit Fallacy that was noted above as one of the two fallacies common to a typical paradox of Zeno. (This omission from Bradis et al.'s (1963) classification scheme is a bit odd, since one of their examples (1963, pp. 19–20) can also be taken as an instance of the Limit Fallacy.) Given that the Limit

Fallacy consists of applying the properties of a sequence's limit to that sequence's individual elements (or vice versa) when that application is not warranted, Misapplication seems to be the logical place to categorize the Limit Fallacy.

Spurious proofs based on this fallacy seem to be common enough to warrant their own subcategory. Besides Zeno's Paradoxes, another common class of spurious proofs resting on the Limit Fallacy is based in continued fractions: two continued fractions are constructed so that, by looking at terms in the numerator and denominator, one can be shown greater than the other, though they compute out to the same value! The trick is that a continued fraction is really a limit of a sequence of standard fractions and, while two sequences may have the same limit, it is perfectly possible for the terms of one to be greater than the corresponding terms of the other if the two sequences converge to the common limit at different rates. (Some other continued-fraction sophisms are better seen as misuses of notation, as the continued-fraction notation conceals information about the elements of the implicit fraction sequence. One such example may be found in Barbeau (2013, p. 137), where the ellipsis conceals the fact that the sequence elements have imaginary terms.)

Misdrawn diagram

Bradis et al.'s (1963, pp. 21–31) final category of spurious proof is that based on "Errors of construction" of a diagram, a category that naturally tends to skew towards geometry. Within this category, they list six subcategories: "Coincident points considered as distinct" (1963, pp. 22–23), "Distinct points considered as coincident" (1963, pp. 23–24), "A point is taken where it cannot lie" (1963, pp. 25–27), "The assumed point of intersection is altogether absent" (1963, pp. 27–29), "A broken line is taken for a straight line" (1963, p. 29), and "A straight line is taken for a broken line" (1963, pp. 29–31). This appears to be a pretty complete listing of the various geometric errors that appear in the literature, although one cannot reject the possibility that other sorts of diagramming errors might be found to constitute the basis for spurious proofs. This list also omits mention of a certain class of geometric proofs that, while they ultimately fall into the above groupings, are numerous enough that they merit their own discussion as an intermediate-level grouping.

Rearranging shapes. That class of problems is the one that involves the rearrangement of shapes; specifically, a set of component shapes is rearranged into two different configurations that seem to have different areas. The apparent difference in area might arise from different results in an area formula (for instance, when the resulting shapes are rectangles with different lengths and heights); or because the resulting shapes appear to have the same perimeter configuration apart from a bump, gap, or interior hole. In either case, the implicit spurious proof is that the areas of the component shapes add up to different sums in the different configurations. The hidden false step is that the equivalence of the two shapes (apart from any bump or gap) is an optical illusion: in one configuration, the component shapes may appear to fit together cleanly but not actually do so; or the dimensions of the two configurations may only appear to match up. This optical illusion ultimately falls into one of Bradis et al.'s (1963) categories: if two shapes do not really fit together, then a straight line is being taken for broken, or vice versa; or if the dimensions of the two figures do not really match, then distinct points are being taken for coincident. However, these rearrangement problems are common enough that they are worth considering as their own intermediate-level category, nested within the Misdrawn-diagram problems, but having the various types of errors nested within them as well.

Rearranging shapes: tangrams. While shape-rearrangement problems exist using all sorts of component shapes, there is a particular subclass that makes use of the standard shapes in a set of tangram tiles. Though it twists the standard terminology somewhat, these puzzles are often referred to as “tangram paradoxes” (Read, 1965, pp. 69–75; Slocum, 2003, p. 100). The tangram tiles have a number of characteristics that lend themselves well to spurious-proof problems, both as a general matter and in the classroom context. This can be seen best by starting with a review of the tangram shapes, as illustrated below in Figure II.*1*.

The standard tangram tile set contains two large isosceles right triangles, one mid-sized isosceles right triangle with hypotenuse equal to a leg of one of the larger triangles, two smaller isosceles right triangles with hypotenuses equal to a leg of the mid-sized triangle, a square with sides equal to the legs of the smaller triangles, and a parallelogram with short sides equal to the legs of the smaller triangles and long sides equal to the hypotenuses of the smaller triangles (or the legs of the mid-sized triangle). The absolute size of the tiles can vary, but if we normalize units such that the legs of the larger triangles equal one, then—based on the Pythagorean Theorem and the basic properties of squares and parallelograms—the other measurements come out as illustrated in Figure II.*1*.

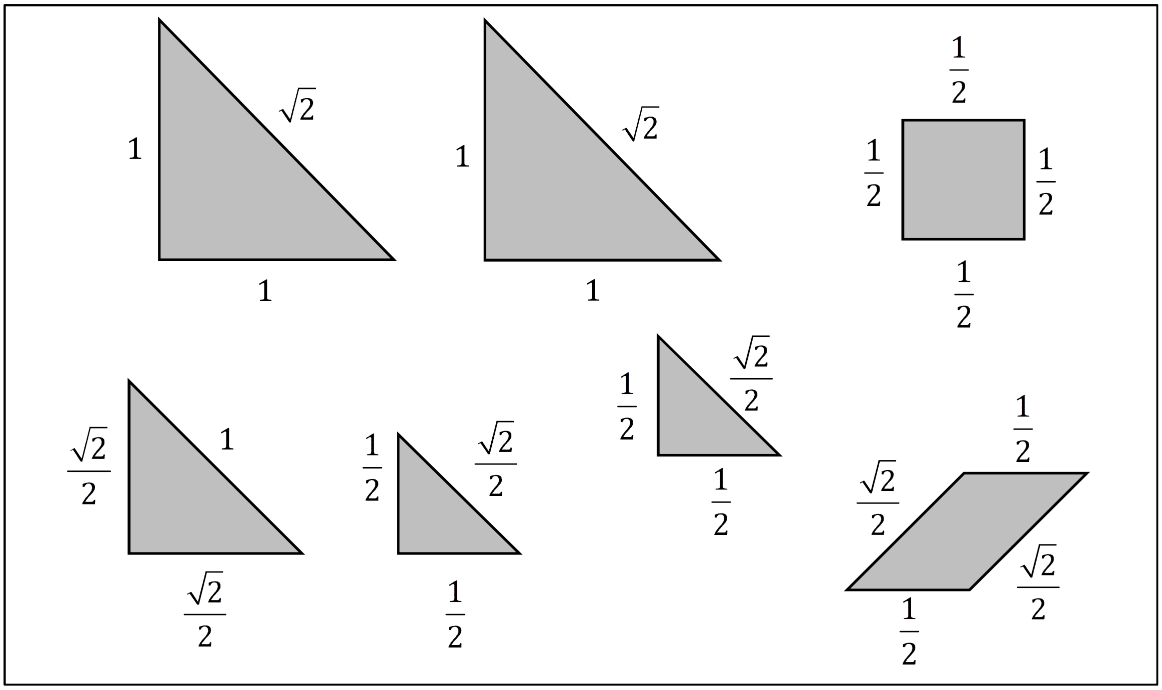


Figure II.1. Shapes and relative sizes of the standard tangram tiles

These measurements suggest the first fact about tangram tiles that makes them well-suited for spurious proofs involving geometric rearrangement: it is easy to arrange tiles so as to create measurements that are very close, but not quite equal. For instance, the hypotenuse of a large triangle measures $\sqrt{2}$, and one could potentially put together any of several segments of length 1 with a segment of length 1/2 (or three segments of length 1/2) in order to create a shape of total length 1.5; that is, similar enough to fool the eye, but not precisely equal.

The question format of a standard tangram puzzle is also fortuitous for spurious-proof problems, in that the necessary deception is easily facilitated. In a standard tangram puzzle, the reader is given the external outline of a shape and the puzzle is to create that outline using the standard tangram tiles. In a tangram paradox, the reader is given two standard tangram puzzles, each consisting of just an external outline. The outlines are similar in most respects, but one possesses a bump or dent or internal gap that implies a difference in area, even though both shapes are to be composed from the same set of tangram tiles. To resolve the paradox, the reader must solve both individual tangram puzzles, and then figure out the subtle difference that leads to the apparent discrepancy in area. However, the format of the basic tangram puzzle, in which the external outline of the final shape is shown and readers must work out the internal arrangement of the tangram tiles on their own, serves here to conceal the false step at the heart of the spurious proof. With the tile arrangements suppressed until the reader solves each of the basic puzzles, it is that much harder to see which edges appear equal in length but really are not.

Tangram paradoxes also have certain particular advantages for classroom use: For one, there are a fair number of problems already published, many dating back to the work of Loyd and Dudeney. For another, tangram manipulatives are commonplace. Though tangram tile sets are typically aimed at younger students than would be working on any spurious-proof exercise, it is not hard to imagine, say, a sixth-grade teacher going to a second-grade teacher to borrow some tiles for a lesson or two. Finally, figuring out the trick typically involves some fairly basic math, at least relative to the usual fare in spurious proofs. Calculating the measurements involved would typically utilize simple application of the Pythagorean Theorem, as well as basic facts about squares and parallelograms, all of which students are likely to be learning in the upper elementary grades (though they wouldn't be learning how to *prove* them until later on). Thus, students could be exposed to this sort of spurious proof fairly early on in their school careers.

Chapter III.

Spurious Proofs as Generators of Cognitive Conflict

One's first intuitive guess at how spurious proofs might be utilized in the math curriculum would likely include two categories: general use as long-form puzzles, and use to teach the proof process. In both these these cases, spurious proofs have indeed had a small presence in the curriculum up to now, but merely as one sort of exercise among many. However, for both these categories, a few strands of recent research have brought up ways in which the distinctive surprise elicited by spurious proofs may give them a usefulness that other sorts of problems do not possess in the same way. Since the value of spurious proofs as long-form problems and in teaching the proof process hinges in both cases on this notion of surprise or "cognitive conflict," these two curriculum uses are discussed here in the same chapter.

General Use as Puzzles

The general sense of present-day math education is that long-form problems requiring some creativity and critical thinking skills are prized over simple exercises in drill and repetition. For instance, the seminal report *Adding It Up* from the National Research Council (2001, p. 335) proclaims that

> the tasks typically assigned to students in many classrooms make only minimal demands on their thinking, relying primarily on memorization or use of procedures without connections to concepts.

> There is growing evidence that students learn best when they are presented with academically challenging work that focuses on sense making and problem solving as well as skill building.

Spurious proofs do constitute one type of long-form problem, but “At the present time, the practice of using paradoxes in mathematical courses is not common” (Kondratieva, 2009b, p. 73). (Kondratieva tends to use the term ‘paradox’ interchangeably with what this paper would designate more strictly as a ‘spurious proof,’ ‘fallacy’ or ‘sophism.’) Some recent authors, though, have suggested that spurious proofs have some unique value as long-form problems due to the fact that their distinctive incongruity stimulates critical thought in a way that more typical problems often do not.

This line of reasoning begins with a general observation from cognitive science that modes of thought tend to be resistant to change: “The case of belief persistence is known to be difficult for pedagogical treatments due to the psychological nature of it: people continue to believe in a claim even after the basis for the claim has been discredited” (Kondratieva, 2009b, p. 65). This is a particular worry in math education, where there is a constant concern that student learning may become excessively routinized: “Many students have developed (over years in high school) a habit of pseudo-learning, mostly in a declarative mode, using explicit rules, given formulas, memorized facts” (Kondratieva, 2009b, p. 68). In a discipline where there is a particular risk of thought patterns becoming mechanistic, the normal human tendency to resist cognitive change may be amplified such that students are deterred from thinking critically and developing understanding of new material.

Spurious proofs offer a potential way to address this issue if their distinctive incongruity can push students to the point where their old thought patterns are simply too contradictory to maintain anymore, and some creative reflection is required to resolve the contradiction and reset their mental framework in a cogent manner. Different authors have used different terms to describe the phenomenon (and some of these terminologies will be discussed below in the next subsection); but Kondratieva (2009b, p. 64), for instance, building on the work of several prior authors, phrases it thusly: "Since a paradox . . . provides a disequilibration, it makes the subject realize the need to re-equilibrate. In this quality, paradoxes are valuable components which stimulate learning and discovery processes by means of restructuring existing schemata of the learner (Rumelhart and Norman, 1978)."

In theory, other forms of mathematical chicanery might serve to provide this sort of disequilibration in a long-form problem; but in practice, the studies that have gone so far as to test their theory with experimental classroom exercises have tended to gravitate toward paradoxes and spurious proofs (Klymchuk, Kachapova, Schott, Sauerbier, & Kachapov, 2011; Kondratieva, 2009b; Mamolo & Zazkis, 2008; Movshovitz-Hadar & Hadass, 1990, 1991; Wessen, 2015) with a certain presence of surprising counterexamples (Klymchuk et al., 2011).

If long-form problems are sought in order to encourage students' creativity, then a spurious proof's particular potential to spark original thought in order to resolve its contradiction would seem to create an unusually valuable sort of long-form problem. One might thus argue that spurious proofs should be utilized in this capacity more than they currently are.

Use to Teach the Proof Process

Another area which would seem to be a prime target for the use of spurious proofs is the teaching the proof process itself. The theory of how to teach proof is not as well developed in the literature as one might like; but, while this means that it is a bit harder to figure out where spurious proofs might fit in to this theory, it also means that finding a place for spurious proofs has a potentially larger contribution to make in this area of the curriculum.

The literature on teaching proof has relatively little in the way of general theory, and only a bit more in the way of particularized suggestions; but after delving in to this literature, we can see that spurious proofs do indeed fit in to what there is of a general theory, as well as into at least one recurring specific suggestion. That particular suggestion is that the need for rigorous proof over less formal justification can be made most apparent to students in cases that surprise or jar the consciousness when students try to address them within a less formal framework. The logic behind this proposition is very similar to the logic of surprise or disequilibrium found in the literature on long-form problems discussed in the previous subsection.

Currently, there is little consensus in the math-education community about how the proof process ought to be taught. There is, though, a fairly extensive discussion on whether it ought to be taught at all. This largely seems to involve math-education theorists like (Ball, Hoyles, Jahnke, & Movshovitz-Hadar, 2002; Fitzgerald, 1996; Gatward, 2011; Hanna, 1989, 1995, 2000a, 2000b; Hanna & Barbeau, 2008; Lutzer, 2005; Markel, 1994; Noormohamed, 2010) arguing for proof in the curriculum while practicing teachers let it phase out. Even an issue of a leading British math-education journal that was supposedly devoted to “Teaching Proof” featured far more articles on whether to teach proof than on how to do it (*Mathematics Teaching*, 1996, no. 155(2)). Perhaps because proof seems to be fading from classrooms below the advanced level, though, there is relatively little literature on the most effective means of teaching it (and thus little that might be inferred about how spurious proofs could fit into that scheme). As one recent survey described the overall situation:

> There is an international recognition of the importance of proof and proving in students’ learning of mathematics at all levels of education, and of the difficulties faced by students and teachers in this area. Also, existing curriculum materials tend to offer inadequate support for classroom work in this area. All of these paint a picture of proof and proving as important but difficult to teach and hard to learn. (A. J. Stylianides & Harel, 2018b, p. v)

A few recent works, though, do try to offer some advice on how to teach the proof process (Ellis, Bieda, & Knuth, 2012; Hanna & de Villiers, 2012; Hanna & Jahnke, 1996; Hanna, Jahnke, & Pulte, 2010; Lin, Hsieh, Hanna, & de Villiers, 2009a, 2009b; A. J. Stylianides & Harel, 2018a). Most of these offer a hodgepodge of specific suggestions rather than a general theory, but one source (D. A. Reid & Knipping, 2010) does collect some suggestions into a general theory. This general theory offers some insight into how spurious proofs might be used to teach the proof process, as does one recurring suggestion: that spurious proofs could be used to jar students’ consciousness into a need to pursue rigorous proofs over informal argumentation.

Spurious Proofs within the General Theory of Teaching Proof

Reid and Knipping (2010, pp. 165–76) survey the existing literature on proof teaching experiments and find four common features of the successful ones: students form conjectures and verify them; proof format is open to the students' devising rather than being preordained; emphasis is placed on the proof process rather than the underlying mathematical content; and classroom norms value explanation beyond what might be formally necessary to establish a proof. For the last two of these features, various sources in the literature suggest a potential role for spurious proofs.

The third feature—emphasis on process rather than content—would clearly be well fulfilled by spurious-proof exercises. This is so, first, because such exercises have a strongly procedural orientation in that the goal is to pick through the proof's steps; and second, because students' focus can be directed in one place rather than two. This is opposed to the historically typical case where "the teaching of proof is always faced with a double problem: that of finding a proof and at the same time of conveying its meaning" (Hanna & Jahnke, 1993).

It takes a little more subtlety to see how spurious proofs are imbued with the fourth of Reid and Knipping's (1996, pp. 165–76) preferred features of proof teaching—an orientation toward explanations. Some authority suggests that secondary reading of a faulty pre-existing proof can help students understand the heuristic process that led the proof's author to his conclusion, as the reader has to sort through what is logically correct about the argument from what is merely appealing to the author's or reader's subconscious (Larvor, 2010) (Polya, 1954a).

The literature (Larvor, 2010, p. 81) (Polya, 1954a) has suggested both pure spurious proofs for this purpose, and proofs that originally began as honest errors—though, if these are consciously presented to later readers in order to spot the error, they are in a sense spurious proofs as well.

The Motivational Shock Value of Spurious Proofs

While it does not show up in Reid and Knipping's (2010, pp. 165–76) survey of the literature, there is another particular suggestion that recurs repeatedly in the proof-education literature and that would seem to offer a ready role for spurious proofs: the proposition that students can sometimes more easily see the need for rigorous proofs instead of less formal arguments when they are presented with problems that are particularly jarring to the consciousness, and for which less formal reasoning clearly no longer suffices. More formally stated,

> Prompting the ultimate move from pragmatic to theoretic knowing requires designing situations so that the pragmatic posture is no longer safe or economical for the learners, while the theoretical posture demonstrates all its advantages. The resultant social and situational challenges are levers which one can use to modify the nature of the learners' commitment to proving (Balacheff, 2010, p. 133).

The need for this tactic arises because

> The difficulty students may have relates not to their lack of mathematical knowledge but to a general human inclination not to question their knowledge and their environment unless there is a tangible contradiction between what is expected after a given action and what is obtained (Balacheff, 2010, p. 123).

This sort of language is rather general, and would seem to admit of many different sorts of mathematical trickery as generators of surprise. Indeed, in some contrast with the literature on cognitive conflict in general problem solving, many different examples of chicanery have been proposed or studied in connection with proof education. These include counterintuitive results (Balacheff, 2010, pp. 118–19, 127–29; de Villiers, 2010, pp. 214–16; Hadas, Hershkowitz, & Schwarz, 2000, pp. 130–39; Movshovitz-Hadar, 1988); surprising counterexamples (de Villiers, 2010, pp. 211–12, 216–20; G. J. Stylianides & Stylianides, 2009); straight problems requiring debate for resolution (Hadas et al., 2000, pp. 139–47; Zaslavsky, 2005); and computer errors (Balacheff, 2010, pp. 120–24); but, of course, spurious proofs have their place as well (Kondratieva, 2009a; Tall et al., 2012, pp. 30–31). All of these articles match theoretical discussion with some sort of empirical treatment: most report on formal studies of actual classroom lessons or sequences of lessons, while Balacheff (2010) and de Villiers (2010) draw on classroom anecdotes.

This shock-value line of reasoning is, of course, very similar to the line of thought discussed in the prior subsection, which backs spurious proofs as general long-form problems because a spurious proof's incongruity stimulates critical thought. Indeed, while the authors in the two lines of research sometimes use different terminology, their ideas are rather similar, and some authors appear in both strands of literature. Many writers have been drawn to the phrase "cognitive conflict" (Lin et al., 2012, pp. 341–42; Movshovitz-Hadar & Hadass, 1990, p. 80, 1991, p. 265; G. J. Stylianides & Stylianides, 2009, p. 319; Zaslavsky, 2005, p. 299; Zaslavsky, Nickerson, Stylianides, Kidron, & Winicki-Landman, 2012, pp. 223–24), while others have sometimes preferred "dis-equilibration" (Kondratieva, 2009b, p. 68) or "surprise" (Movshovitz-Hadar, 1988, p. 14). However, there seem to be a great many theories in this area that seem to be very similar and very closely interrelated. One paper's authors, for instance, noted that

> The terms "cognitive conflict" and "uncertainty" have overlapping but distinctive meanings (Zaslavsky 2005) [sic]; here, we use them interchangeably: We also subsume in these terms a range of other related terms from the mathematics education literature, such as contradiction, perplexity, and surprise. Zaslavsky (2005) provided a detailed discussion about the roots of the notion of cognitive conflict in Dewey's concept of reflective thinking and its relations to psychological theories such as Piaget's equilibration theory, Festinger's theory of cognitive dissonance, and Berlyne's theory of conceptual conflict (Zaslavsky et al., 2012, p. 223).

Even when authors use different terms, their ideas may come from the same source. Zaslavsky (2005, p. 299) for instance, in the proof-education literature, utilizes the term "cognitive conflict"; while Kondratieva (2009b, pp. 63–64), writing about general problem-solving, prefers "dis-equilibration"; but both cite Piaget as a source for their theories.

The modern strand of literature on 'cognitive conflict,' or whatever one wishes to call it, seems to trace back to Movshovitz-Hadar (1988). Some authors, though, trace the precursor ideas back to Dewey (1933) (cited in Zaslavsky, 2005, p. 299) or Polya (Polya, 1954a) (cited in Larvor, 2010, p. 81). Movshovitz-Hadar and Hadass (1991, p. 80) even trace the principle back to the Book of Exodus, where Moses perceives an apparent contradiction—"The bush burned with fire, and the bush was not consumed" (Exodus 3:2)—and then must take action to resolve the cognitive conflict: "I will now turn aside, and see this great sight, why the bush is not burnt" (Exodus 3:3).

In sum, then, cognitive conflict and similar concepts have been proposed in the math-education literature as giving spurious proofs, with their uncommon surprise value, a certain distinctive utility both as general long-form problems and as instruments of teaching the proof process. The latter value is augmented as well by their potential use in exercises where proofs are analyzed secondarily rather than being composed from scratch.

Chapter IV.

Spurious Proofs as Builders of Productive Disposition

In the last case, reinforcing perceptions of the coherency of math, these perceptions have been recognized in the literature for some time as a key component of proficiency in mathematics; but the literature has also noted that this component of proficiency has received less attention than other components that more directly relate to the performance of mathematics. Part of the reason for this gap in the literature would seem to be the apparent difficulty of finding mathematical activities that are particularly targeted toward reinforcing perceptions of math's coherency, and spurious proofs would seem to be one of the few sorts of problems that are specifically tailored toward this end. The math-education literature also suggests, however, that spurious proofs have an untapped potential as an exercise to reinforce one of the core components of mathematical proficiency. In theory, at least, the most distinctive and noteworthy educational use of spurious proofs may come from their ability to reinforce students' perceptions of the mathematical system as a cogent and worthwhile object of study.

The importance of this perception in developing mathematical proficiency was first articulated by the report *Adding It Up: Helping Children Learn Mathematics* (National Research Council, 2001). Since its publication, this report has become an important part of mathematical-education literature, with its framework of analysis guiding many subsequent efforts to understand and improve mathematics education. (This is my general impression from reviewing the literature, though a more rigorous basis for the claim would be worthwhile in the final thesis.) The core of the report's framework of analysis is that overall mathematical proficiency comprises five components or strands:

> • *conceptual understanding*—comprehension of mathematical concepts, operations, and relations
> • *procedural fluency*—skill in carrying out procedures flexibly, accurately, efficiently, and appropriately
> • *strategic competence*—ability to formulate, represent, and solve mathematical problems
> • *adaptive reasoning*—capacity for logical thought, reflection, explanation, and justification
> • *productive disposition*—habitual inclination to see mathematics as sensible, useful, and worthwhile, coupled with a belief in diligence and one's own efficacy (National Research Council, 2001, p. 116) (the list is a direct quote; bulleting and emphasis original)

The fifth strand, productive disposition, is something of an odd duck relative to the others in both its substance and its treatment in the math-education literature. Substantively, as one notes readily from the above definitions, all the other strands relate in one way or another to the direct performance of mathematical tasks; and, in principle, this means that they can fairly readily be developed using exercises catered to those tasks. In the literature, productive disposition has been described as "the least researched strand of mathematical proficiency" (Siegfried, 2012, p. 5). These two facts together led J.M. Siegfried to dub productive disposition the "Hidden Strand" in the very title of his (2012) dissertation: it has been 'hidden' from the literature, and "People use their productive dispositions to help them solve problems and build new mathematical knowledge, but their dispositions generally are not evident in their written work" (Siegfried, 2012, p. 22).

These facts also mean that the present thesis has the potential to contribute knowledge both by generally filling in a sparse literature, and by offering spurious proofs as a category of mathematical problems that can be specifically tailored to developing productive dispositions. After a bit of additional background on the nature and cultivation of the various strands of proficiency, and on the place of productive disposition in the math-education literature, it is easy to see as the next step a role for spurious proofs in the formation of a productive disposition.

The Strands of Proficiency

As noted above, the framework of the five interwoven strands of proficiency was first crystallized in the report *Adding It Up: Helping Children Learn Mathematics* (National Research Council, 2001), although the basic ideas in that framework had been percolating in other literature (National Research Council, 2001, pp. 117–18). Since then, the framework has continued to be a widely used guide in studying and assessing mathematics education It has been incorporated into the Common Core State Standards Initiative (Grady, 2016, p. 516), and the Web of Science ("Web of Science," 2018) database indicates more than 700 citations over the years. Looking into each of the strands in a little more detail, one can see that productive disposition is unique among the strands in not being directly related to the conduct of mathematics, and therefore presents some distinct challenges for teaching.

Conceptual understanding. *Adding It Up* denotes conceptual understanding as fundamentally the ability to link bits of mathematical knowledge together, so that students "have organized their knowledge into a coherent whole" (2001, p. 118). This, in turn, allows students to learn new material more readily by linking it to what they already know, and to remember their mathematical knowledge more fully (2001, pp. 118–20). A key indicator of conceptual understanding is the ability to move between different representations of the same phenomenon (2001, p. 119).

Procedural fluency. According to *Adding It Up*, procedural fluency "refers to knowledge of procedures, knowledge of when and how to use them appropriately, and skill in performing them flexibly, accurately, and efficiently" (2001, p. 121). In other words, it includes the basic knowledge of common solution techniques that is the subject of so much 'drill and kill' consternation, but also a more in-depth knowledge of those techniques that allows students to be confident they are using a given technique in the right situation and adapting it properly if necessary for a particular context. While a naïve view might tend to associate procedural fluency with algorithm memorization and conceptual understanding with higher-level knowledge, in fact the two strands are interrelated since proper use of procedures involves some higher-level knowledge, and since conceptual understanding doesn't arise without a certain amount of practice in basic procedures (2001, p. 122). A typical indicator of strong procedural fluency is skill at mental math by adapting standard techniques in a way that can be computed quickly and paperlessly given the problem at hand (2001, pp. 121–22).

Strategic competence. Strategic competence comprises "problem solving and problem formulation" (National Research Council, 2001, p. 124), meaning both the ability to address mathematical questions that are more complicated than a standardized technique can handle, and the ability to reduce real-life situations to a mathematical representation that can be addressed either by a standardized or more complicated technique. Strategic competence depends on both conceptual understanding and procedural fluency, since formulating and solving a problem depends on seeing how all the concepts at play in the problem relate to each other, as well as on knowing which procedures are likely to be handy at various stages along the way, and on being able to resolve those intermediate stages as routine procedures (2001, p. 127). In turn, strategic competence helps to build procedural fluency as solving more and more problems gives students more techniques to integrate into their procedural toolkit (2001, pp. 128–29).

Adaptive reasoning. The strand of adaptive reasoning is defined as "the capacity to think logically about the relationships among concepts and situations" in mathematics (National Research Council, 2001, p. 129), with its fundamental expression being "the ability to justify one's work" either by formal proof or by less rigorous argumentation (2001, p. 130). This strand is intertwined heavily with the prior three: making a sound mathematical argument depends on understanding the concepts involved, on executing intermediate procedures, and on having a good strategic approach to the argument (2001, pp. 130–31); at the same time, arguments through adaptive reasoning validate procedures and solution techniques, and reinforce conceptual understanding (2001, p. 129).

Productive disposition. Productive disposition denotes "the tendency to see sense in mathematics, to perceive it as both useful and worthwhile, to believe that steady effort in learning mathematics pays off, and to see oneself as an effective learner and doer of mathematics" (National Research Council, 2001, p. 131). For obvious reasons, this strand of proficiency is tightly interwoven with the others: Viewing math as sensible and oneself as competent at it will tend to make students put more effort into all the practical aspects of doing mathematics; but, at the same time, having success at actually solving problems will tend to reinforce students' view of math as cogent and sensible, as well as students' views of their own competence. As *Adding It Up* puts it, "A productive disposition develops when the other strands do and helps each of them develop" (2001, p. 131).

The reader may have noted here that the strand of productive disposition in fact appears to comprise two substrands, with a level of intertwining that is not easily determined at first glance. The first substrand comprises students' attitudes and beliefs about mathematics—in the definition quoted above, "the tendency to see sense in mathematics" and "to perceive it as both useful and worthwhile" (National Research Council, 2001, p. 131). The second substrand comprises students' attitudes and beliefs about themselves—"the tendency . . . to see oneself as an effective learner and doer of mathematics" (National Research Council, 2001, p. 131). The first substrand is the one that is more pertinent to the present study, as we are investigating whether a particular sort of mathematical exercise can help build a productive disposition. The second substrand, while important, is best addressed by other sorts of programs or interventions than those being studied here.

Adding It Up seems to maintain that there is some interaction between those two substrands. The third clause of its definition, for instance, "the tendency . . . to believe that steady effort in learning mathematics pays off" (National Research Council, 2001, p. 131) would logically only hold true when a student believes both that the underlying math has a sense to it that can be grasped with effort, and that the student is skilled enough that the effort at grasping the math is not wasted.

However, the topic of the respective substrands is rarely discussed directly in the literature, leaving many open questions about the relative magnitudes of the two strands, and about the level of their interaction. Put another way, it is not immediately obvious from the literature how much a productive disposition is, or can be, developed from students engaging with the mathematics; and how much it is developed from teachers reinforcing students' self-esteem and self-efficacy beliefs through direct interaction with the students, in a manner akin to 'cheerleading.'

When one examines the literature closely, one finds that there is a fair amount of work both directly and indirectly supporting the importance of the self-esteem aspects of productive disposition; but there are also a number of studies implying that student interactions with mathematics have an important role, and that the interaction between the self-esteem and mathematical substrands is deep. This literature is addressed in the next subsection.

Productive Disposition in the Literature

As noted above at p. 42, productive disposition gets relatively little research attention relative to the other strands of proficiency. This is especially true regarding research that explicitly cites the five-strand framework of *Adding It Up*, but somewhat less so if one considers research into closely related concepts that particular authors happen to designate with different labels, such as 'student attitudes' or 'beliefs' about math. Beyond the one major study of the productive-disposition as such, the literature on particular aspects of productive disposition (or closely related concepts) tends to fall into three main categories: self-efficacy beliefs related to demography, productive disposition and mathematical beliefs or attitudes among current and prospective teachers, and classroom practices that promote productive disposition. Those three bodies of literature are each addressed in respective subsections a bit further below, but first it is worth looking at the one major study that addresses the whole matter of productive disposition as its central topic.

That major study is John Michael Siegfried's (2012) dissertation *The Hidden Strand of Mathematical Proficiency: Defining and Assessing for Productive Disposition in Elementary School Teachers' Mathematical Content Knowledge*. As the full title indicates, Siegfried's concern was less with what sort of classroom practices can promote a productive disposition; and more with what sort of behaviors are indicative of a productive disposition, and how can teachers observe and assess for them. As a result, the dissertation does not directly bear on the topic at issue here, but it does contain a worthwhile overview of the characteristics of productive disposition (2012, pp. 20–52).

It also sheds some additional light on what constitutes a productive disposition through its analysis of particular student behaviors that correlate with it. Siegfried looked at the definition of productive disposition posited by *Adding It Up* and found eight concepts in the math-education literature that were closely related to that definition:

- *Affect*—a person‘s feelings and attitudes that shape the way one looks at the world

- *Beliefs*—psychological understandings about how one perceives the world to be
- *Goals*—the states that human beings desire to obtain
- Identity—qualities people recognize in themselves or that are recognized by others (a person's type or kind)
- *Mathematical Integrity*—knowing what one knows, knowing what was does not know, and being honest about these assessments
- *Motivation*—the inclination people have to do certain things and avoid doing others
- *Risk Taking*—willingness to ask questions or share ideas that may expose one's misconceptions or weaknesses
- *Self-efficacy*—a person's own belief in his or her ability to take action on a particular problematic situation (Siegfried, 2012, p. 24) (the list is a direct quote; bulleting and emphasis original)

He then combined the above list with the four elements of the definition of productive disposition from *Adding It Up* (2012, pp. 59–62) so as to form a somewhat broader and differently categorized definition of productive disposition. In turn, from the math-education literature and from a preliminary empirical analysis, he identified sets of behaviors that were characteristic of the various categories of his broadly-defined productive disposition. These behaviors are listed in Table IV.1 below.

Not all of the categories and behaviors in Table IV.1 are directly relevant to the present inquiry, but they do give some additional illustration of what is constitutes a productive disposition. They are therefore worth keeping in mind when considering the various bodies of literature dealing with that strand of proficiency.

Table IV.1

Siegfried's (2012) Indicia of Productive Disposition

Potential Categories	Evidence
1. Mathematics as a sense-making endeavor	a. Tries to make sense of the task b. Considers alternative approaches c. Asks if answer seems logical d. Is troubled by inconsistencies
2. Mathematics as beautiful or useful and worthwhile	a. Shows interest in the task through engagement b. Shows interest in the task in comments about the task c. Shows a sense of wonder
3. Beliefs that one can, with appropriate effort, learn mathematics	a. Believes making progress on the task is doable b. Persists c. Does not avoid frustration
4. Mathematical habits of mind	a. Asks questions about the mathematics or about an approach (one's own or another's) b. Shows appreciation for one's solution c. Seeks and provides clarifications
5. Mathematical integrity and academic risk taking	a. Has a sense for when one has completed a task (whether or not one continues) b. Is willing to question one's self c. Is willing to offer tentative ideas d. Recognizes worthy and unworthy confusions
6. Positive goals and motivation	a. Defines progress as learning through grappling, not just getting an answer b. Shows pleasure or excitement about a particular way of reasoning c. Engages longer or willingly reengages with difficult tasks d. Interprets frustration, when experienced, as a natural component of problem solving and not as a statement of one's mathematical competence
7. Self-Efficacy	a. Seems confident in one's own abilities and skills for solving the task b. Seems confident in one's knowledge

Note. Table text and organization are reproduced from Siegfried (2012, p. 109). Table format is altered, and emphasis related to Siegfried's research process is suppressed.

Demography. One line of literature on productive disposition centers on demographic categories, specifically on the phenomenon that students from a given demographic group may tend to feel—or lack—a group-based confidence that someone from their particular demographic category is likely to be able to perform well at mathematics.

One way in which this has been observed to happen is ‘stereotype threat,’ in which “good students who care about their performance in mathematics and who belong to groups stereotyped as being poor at mathematics perform poorly on difficult mathematics problems under conditions in which they feel pressure to conform to the stereotype” (National Research Council, 2001, p. 133).

Another notable aspect of this body of literature is its partial overlap with “growth mindset” theory. As a general proposition, growth mindset theory holds that students tend to see mathematical ability either as something learnable or as something innately fixed, with the former being more conducive to the learning process (Hocker, 2017, pp. 4–5). *Adding It Up* associates this idea with productive disposition, citing growth-mindset guru Carol Dweck (1986) for the proposition that “Students who view their mathematical ability as fixed and test questions as measuring their ability rather than providing opportunities to learn are likely to avoid challenging problems and be easily discouraged by failure” (National Research Council, 2001, pp. 132–33). A demographic component enters this line of thought because fixed-mindset behavior is observed more frequently in girls (Dweck, 1986, pp. 1043–45).

The literature in this line is fairly extensive, sometimes even percolating into the popular press (Eveleth, 2013; Paul, 2012). While *Adding It Up* links this area of study to productive disposition, studies about group-based math attitudes or stereotype threat do not always link their concepts back to the *Adding It Up* framework. For instance, *Adding It Up* (2001, pp. 132–33) cites the growth-mindset work of Dweck (1986) as being related to productive disposition; but some recent growth-mindset studies cite extensively to various works of Dweck, and a great deal of other growth-mindset literature, while omitting all mention of the strands-of-proficiency framework (Hocker, 2017; Jones, 2016).

<u>Teacher attitudes</u>. Another collection of studies addresses productive disposition amongst current and aspiring teachers.

Several studies consider ways of assessing productive disposition in teachers (E. Jacobson & Kilpatrick, 2015; Jong & Hodges, 2015; Lewis, Fischman, & Riggs, 2015; Philipp & Siegfried, 2015; Siegfried, 2012).

Others focus on means of improving productive disposition in teachers; but within this group, a dichotomy should be noted: There are a few studies analyzing short-term classroom practices that purport to address productive disposition, but where the practice happens to have been tested among teachers in a pre-service or continuing-education context, rather than in an ordinary classroom. While formally involving teachers, these studies are a better conceptual fit with general studies of classroom practice, and so are analyzed in the subsection below. On the other hand, several studies examine potential methods of enhancing the long-term teacher-education curriculum in order to enhance new teachers' productive dispositions (E.

D. Jacobson, 2017; Philippou & Christou, 1998; Schram, Wilcox, Lappan, & Lanier, 1988; Swars, Smith, Smith, & Hart, 2009). While worth noting, these studies do not necessarily say much about how to instill a productive disposition in younger students. For one thing, as A. J. Stylianides & Stylianides (2014, p. 10) note, "their long duration suggests that it is hard both to replicate their success in other settings and to incorporate them into existing educational programs." For another, prospective teachers are obviously at far different stages of intellectual and personal development from younger students, and it is hardly clear than a practice tailored to the teacher-education curriculum would translate to a typical classroom setting.

Classroom practices. A number of studies do look, however, at short-term practices which purport to be helpful in building productive dispositions in a more typical classroom. Some of these are fairly obvious, and others less so.

Among the more obvious ones are—implicitly—all the studies of the previous subsection, for the straightforward reason that teachers with productive dispositions will tend to produce students with productive dispositions as well. That attitudes and beliefs can be contagious is hardly surprising; but researchers have, indeed, confirmed the obvious:

> One's self-efficacy beliefs can also develop through inferences drawn by the individual from social comparison: "[s]eeing others perform threatening activities without adverse consequences can generate expectations in observers that they too will improve if they intensify and persist in their efforts" (Bandura, 1977, p. 197). An observer's self-efficacy beliefs are more likely to increase if the observer can identify himself/herself with individuals who succeed and if the success can be attributed to effort and perseverance. (A. J. Stylianides & Stylianides, 2014, p. 11)

Another seemingly obvious technique for building productive disposition is to emphasize classroom activities involving long-form problems and creative reasoning. The 'null hypothesis' held by many students is that "mathematics is mostly a set of rules and that learning mathematics means memorizing the rules" (National Research Council, 2001, p. 141). A classroom environment with more active problem solving and less memorization would seem, then, to be a clear way to induce a 'rejection' of this 'null hypothesis' (Boaler, 1998; Frank, 1988; Garofalo, 1989; Grady, 2016; Stodolsky, 1985).

Finally, a number of authors posit more particularized, short-term classroom interventions that show promise at building productive disposition (Blazar & Kraft, 2016; Hocker, 2017; Sriraman, 2009; A. J. Stylianides & Stylianides, 2014). It is worth noting about Hocker (2017) that hers may not be the only study from the growth-mindset school to offer up a classroom practice that builds productive disposition; with the observed tendency of those authors to be relatively isolated in their silo, other such studies may exist while not being readily apparent from the productive-disposition literature.

The other three studies in this category (Blazar & Kraft, 2016; Sriraman, 2009; A. J. Stylianides & Stylianides, 2014) are each discussed in more detail in the next subsection, where spurious proofs are analyzed for their own fitness as classroom practices promoting productive disposition.

Using Spurious Proofs to Develop Productive Disposition

In various ways, several of the arguments in this literature about encouraging productive disposition have implications that suggest a possible role for spurious proofs in this process. However, no one to date has examined the most obvious role for spurious proofs in this process: that a spurious proof challenges the coherency of the mathematical system—the core of the mathematical aspect of productive disposition—and asks the student to resolve that challenge and re-establish the system's coherency in an active problem-solving venture. After looking at some previous suggestions for promoting productive disposition, and their relevance to spurious proofs, that core relationship is examined in a little more detail.

One threshold question that confronts the issue of what classroom practices can promote productive disposition is the question of the relative importance of the mathematical and personal aspects of productive disposition; and here, a particular study by Blazar and Kraft (2016) suggests that the personal aspects ought not be overemphasized to the exclusion of the mathematical. Recall that productive disposition includes both students' beliefs about mathematics and students' beliefs about themselves as doers of math. Thus, there is potentially an open question of how much productive disposition develops from interacting with mathematics and how much it is simply self-esteem building based on interactions with the teacher. Blazar and Kraft (2016) conducted a statistical study to relate a set of student outcome variables, both academic and non-academic, to a set of independent variables including various observed teaching practices as well as background variables for schools, teachers, and students; and fixed effects for schools, teachers, and district-years. A very telling result in the present context is that in their regression for students' self-reported self-efficacy, Blazar and Kraft (2016, p. 157) found (among other effects) not just a significant positive effect for teacher emotional support, as one would expect given that self-efficacy is a personal aspect of productive disposition; but also a significant negative effect for mathematical errors. In other words, when analyzing an aspect of productive disposition that is primarily personal, they found a significant effect from a phenomenon that is primarily mathematical—the presentation of math by the teacher that is not, in fact, sensible and coherent.

(It should be noted that one potential objection to this result is mooted by Blazar and Kraft's (2016) statistical design. Conceivably, the negative correlation between teacher errors and student self-efficacy might be explained because error-prone teachers have lackadaisical attitudes about math, and these errors might be transmitted directly while spuriously being picked up by the teacher-error variable. However, since Blazar and Kraft (2016, pp. 152–53) included teacher and teacher-class dummy variables, general teacher attitudes should be picked up there, rather than through errors.)

It is thus hard to argue with Blazar and Kraft's (2016) finding that teacher mathematical errors have a negative impact on student self-efficacy beliefs. This result is important because it implies that the math matters. Building a productive disposition, even in aspects that are primarily personal, cannot be done most effectively by simple cheerleading. Instead, it also requires intervention in the form of experiencing and doing mathematics. The question then is: just what sorts of interventions can help promote productive disposition?

Among prior suggestions for promoting productive disposition, the first one to point out is the common one to utilize more creative problem solving. As noted above at pp. 31–34, spurious proofs have a particular value as long-form problems because their incongruity creates a cognitive conflict that leads to an exertion of critical thinking in order to resolve the conflict. If a teacher is simply looking for examples of problems that go beyond rote memorization, spurious proofs have a special utility as that sort of problem.

Turning our attention to more particular examples of promoting productive disposition, A. J. Stylianides & Stylianides (2014) attempt to create a short-term means of boosting productive disposition through performing one mathematical problem-solving exercise. They do develop one that shows promise, but they also come up with four design principles for creating future problems to serve this purpose. The principles were derived by building on psychological theories of how beliefs become ingrained into memory, and they are:

> 1. The problem has memorable characteristics (e.g., name, context)
> 2. The problem initially seems to be unsolvable
> 3. The problem includes few clearly identifiable mathematical referents (numbers or formulas) that by themselves offer insufficient information for its solution
> 4. A solution to the problem
>   (a) is within students' capabilities after perseverance (and support from peers or limited instructor scaffolding)
>   (b) involves achievement of several milestones (2014, p. 11) (list is a direct quote; numbering original)

All of these would seem to be well served by spurious proofs, either directly or at least in general spirit: Spurious proofs are generally memorable with the apparent contradictions they present. They also initially seem to be unsolvable if their false step is well hidden. They do sometimes include fair amounts of numbers and formulas, but they should still tend to satisfy the spirit of principle 3, since a spurious-proof problem is not meant to be solved by calculating from the numbers and formulas: the logic of principle 3 is that "[t]he fact that the students will need to consider other kinds of referents [than numbers or formulas] to solve the problem is expected to help expand their view about the nature or importance of different kinds of referents in mathematical problems" (2014, p. 12); and with a spurious proof, the referents needed to find the fallacy are the logical arguments behind each step, not the particular content of any formulas in the proof. Principle 4(a) is readily satisfied as long as the proof chosen is of an appropriate ability level for the students doing the exercise. Principle 4(b) is not necessarily satisfied directly, but like with principle 3, a spurious proof might easily enough satisfy the same logical criteria: the rationale behind principle 4(b) is that "[h]ad the solution involved achievement of only one milestone (i.e., if one key idea or step sufficed for the solution), then once a student in a small group had an insight into how to achieve that milestone the rest of the group would be deprived of opportunities to" contribute (2014, p. 12); and, while spurious proofs might often need just one key idea to resolve, they are more amenable than most math problems to separating into ex ante subtasks (say, by assigning one step to each person in a group to analyze as a potential fallacy) so that each student can have a role in the solution.

A final study is worth mentioning here because it utilized a mathematical paradox—a close cousin of spurious proofs—in an attempt to mold the mathematical beliefs of pre-service teachers (Sriraman, 2009). Sriraman (2009) found a positive effect from an exercise using Russell's Paradox in an effort to foster more creative and nuanced beliefs about mathematics. However, his results are a bit difficult to interpret here because he utilized a different analytical framework for mathematical beliefs from what is set out in *Adding It Up*. Citing P. Ernest, Sriraman (2009) divided mathematical beliefs into

> the instrumentalist view, the Platonist view, and the problem solving view. The instrumentalist sees mathematics as a collection of facts and procedures which have utility. The Platonist sees mathematics as a static but unified body of knowledge. Mathematics is discovered, not created. The problem solving view looks on mathematics as continually expanding and yet lacking ontological certainty. . . . Ernest also describes absolutist and fallibilist views of mathematical certainty. The absolutist sees mathematics as completely certain and the fallibilist recognizes that mathematical truth may be challenged and revised (Ernest, 1991, p. 3) (the portion of the quote prior to the ellipsis refers to a different work of Ernest).

Though Sriraman doesn't spell this out explicitly, an instrumental outlook seems to be generally aligned with absolutism; a problem-solving outlook with fallibilism; and a Platonist outlook potentially with either one, with a fallibilist-Platonist view reflecting a belief that underlying mathematical truth is absolute but that humans are sometimes fallible at grasping it.

Among the prospective teachers working with Russell's Paradox, Sriraman (2009, p. 90) rigorously documented a set of behaviors that other literature had associated with polymathy, or the ability to move analytically between academic disciplines or skill groups. Combining that with a qualitative overview of student responses in which the participants reported having a generally more philosophical, literary, or artistic outlook on mathematics after the exercise, Sriraman (2009, pp. 84–89) inferred that the student teachers had tended to move their belief systems from an instrumental or absolutist-Platonist orientation to a problem-solving or fallibilist-Platonist one, and had thus become more comfortable in the previously somewhat alien mathematical discipline.

However, when attempting to translate these results to the productive-disposition framework, it is not entirely clear what Sriraman was picking up. Specifically, since his instrumental–Platonist–problem-solving belief framework does not have the same division into mathematical and personal beliefs that productive disposition has, it is uncertain whether Sriraman's subjects were developing a more structured and integrated view of mathematics itself; or were finding that they, as people with generally more literary-philosophical orientations, were becoming more comfortable with math than they were before. It is worth keeping in mind, though, that Sriraman's analytical framework derives from a literature that sees paradoxes as gateways to polymathy across all disciplines, not just mathematics (2009, pp. 78–79). Since not all disciplines face the same issue math does of fighting a common perception as an arbitrary collection of rules and facts to be memorized, this is some indication that paradoxes are doing more than offering a haven for a literary mindset. To have the same effect across many disciplines, they more likely must have a common effect of making new material seem sensible.

Thus, a number of studies have found results that point to a role for spurious proofs in building productive disposition. Previous analyses have found that classroom practices focused on math, not just on personal beliefs, are critical; that math exercises can build productive disposition, and that the criteria for exercises to do so are favorable to spurious proofs; and even that the closely related category of paradoxes is specifically helpful in building a productive belief system in the area of math. However, all these studies seem to miss the most direct potential impact for spurious proofs on productive disposition: that a spurious proof is an active problem that is aimed directly at the mathematical aspects of productive disposition.

With a spurious-proof problem, students are confronted with a proposition that seems to contradict the mathematical knowledge that they have learned before—and thus the very premise that mathematics is non-contradictory. Then, they must go through the active process of finding the false step in the proof, demonstrating that the seemingly impossible proposition really is impossible like it is supposed to be, and that the mathematical system is coherent after all. Since the coherency of the mathematical system is central to the mathematical aspects of productive disposition, spurious proofs seem to be one of the few types of problems—perhaps the only one—where a productive disposition is reinforced through an active problem-solving process rather than passively relied on in the background. This is a particular case of the common educational nostrum that "students learn best when they are presented with academically challenging work that focuses on sense making and problem solving as well as skill building" (National Research Council, 2001, p. 335); but here, the need for active problems to solve is even more acute because the particular topic—the sensibility of the mathematical system—is so rarely explicit in any problem at all.

Unfortunately, studies to date seem to have missed the opportunity to analyze this potential value for spurious proofs. The closest approach to it might be

Movshovitz-Hadar and Hadass (1990, pp. 279–80, 1991, pp. 85–86), who describe the role of a spurious proof in breaking down "Knowledge fragility" (knowledge that is incomplete and poorly integrated) and resetting it on a firmer basis.

Nonetheless, there is substantial evidence that spurious proofs have a large and underappreciated role to play in teaching mathematics: Their stimulation of cognitive conflict gives them a particular value as long-form problems. The same feature makes them valuable in teaching the proof process, as does the opportunity they present to make a secondary analysis of an existing proof. The mathematical aspects of productive disposition cannot be ignored in favor of the personal; and in various ways, problems with similar features to spurious proofs have been shown to boost productive disposition in this way. It should not surprise us, then, that spurious proofs could have a direct role to play in developing this aspect of proficiency that heretofore has received too little attention. With the goal of suggesting how this might be done, the next chapter presents several possible classroom exercises utilizing spurious proofs.

Chapter V.

Classroom Exercises

So far, this thesis has discussed the value of traditional spurious proofs in a traditional context—suitable to be worked on by individual students as homework or class exercises. However, it is also worth contrasting this sort of use with ways in which spurious proofs can underpin more elaborate learning activities. This can be done by illustrating a relatively standard spurious-proof problem based on a tangram paradox exercise, alongside a more elaborate lesson consisting of the composition of an original spurious proof based on one of Zeno's paradoxes.

A Tangram Paradox

As a spurious-proof exercise, the following tangram paradox follows the usual format fairly closely: demonstrate the apparent contradiction, then guide the class toward an understanding of why the illusion occurs. There are, however, possibilities for adding some bells and whistles.

The underlying puzzle is, so far as can be determined by appropriate due diligence, original to the present author. It was devised independently, and does not appear in any of the usual sources for tangram paradoxes (S. Loyd, 1903, 1914, 1959, 1960); (Dudeney, 1908, 1917, 1967); (Goodrich, 1817, 1818); (Voronets, 1930); (Movshovitz-Hadar & Webb, 1998; Nufer, 2017; Read, 1965; Sarcone, 2017; Van Note, 1973); (Gardner, 1987, 1988b). It should be noted, though, that very similar paradoxes founded on the same basic trick are found in S. Loyd (1903, nos. 25–26) and Sarcone (2017), and a standard puzzle very similar to the target shape of the present paradox but featuring an hourglass-shaped hole may be found in S. Loyd (1903, pp. 3, 36), Goodrich (1817, 1818) and Gardner (1988b, p. 33). It is particularly worth noting that the present puzzle does not appear in Slocum (2003), which attempts an extensive, though not absolutely complete, cataloging of historical tangram puzzles. Slocum (2003, p. 100) does, however, reprint several paradoxes based on the same trick, and make the paradoxical connection for the hourglass-hole figure noted above. Finally, it should be noted that, given the historically informal customs for attribution and citation in the realm of recreational mathematics, it is impossible to say for certain that the puzzle has not been devised before, but simply has not found its way into the more formal literature.

The exercise may be done either by the whole class together or in smaller discussion groups, and assuming any necessary background is out of the way—for instance, the class is familiar with tangram tiles, and students have sets of them at hand (though cutting out their own tiles could be a bit of a diverting problem in itself), the lesson proceeds as follows:

The teacher presents the initial base shape

This is a perfect square, and the students should complete this fairly readily given that this is a standard way of arranging and storing the tiles. Allowing for rotation, the standard solution is:

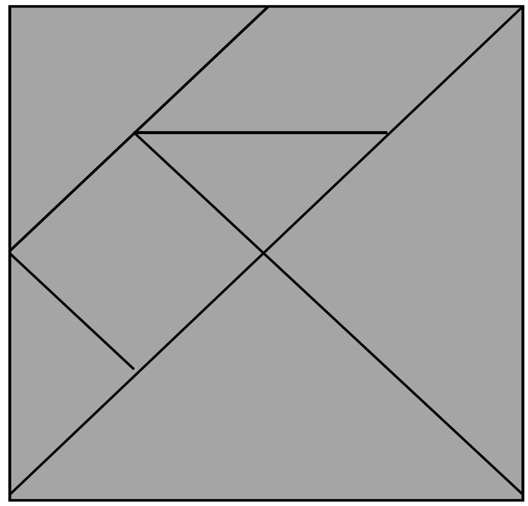


Figure V.1. Tangram tiles forming a square.

The teacher presents the target of the paradox

Next, the teacher presents the target shape that generates the paradox, which is:

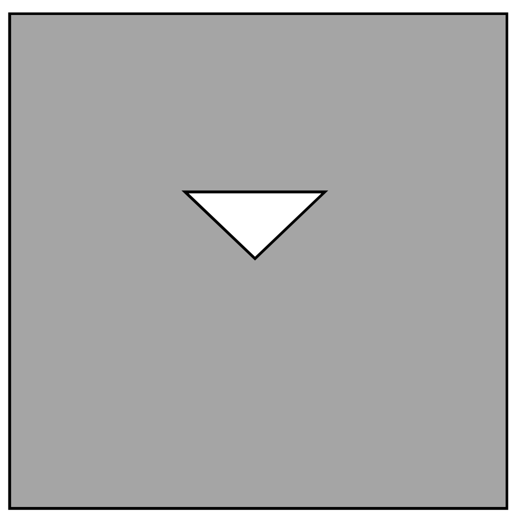


Figure V.2. Target shape of tangram paradox

Students work out the solution

Potentially, students could work on the problem individually as seatwork, in small discussion groups, or together as an entire class. There may not be much potential for hints that fall short of giving away the solution, but the teacher should be prepared to offer them if the occasion arises. Eventually, the class should arrive at the following solution:

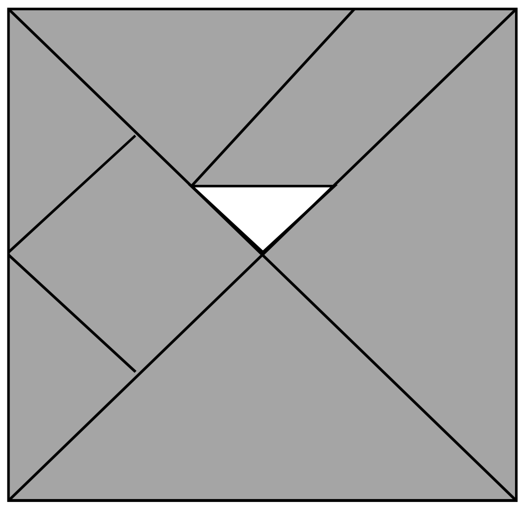


Figure V.3. Solution to the target tangram paradox

Closing discussion

After the solution has been reached (and presented to the class, if the whole class has not been working together up to that point), the teacher should prompt the class, if they have not done so already, to recognize the paradox: that the same set of component tangram tiles appears to produce shapes that add up to different areas—because they seem to have the same exterior outline, but one has the internal gap that would seem to imply a diminished area.

With such hints as may be necessary, the class should be led to see where the fallacy is: that the top edge of the 'square' in Figure V.*3* is actually a bit longer than the other sides. If one uses the normalized units from Figure II.*1*, then the top edge measures 1½ while each of the other edges measures $\sqrt{2}$. The upper trapezoid thus does not really fit cleanly into the rest of the square, which may be apparent from a close look at students' tile arrangements.

Some further discussion ought to be done in order to establish why this is so. That is, the class ought to resort to the Pythagorean Theorem and the basic properties of triangles, squares, and parallelograms in order to derive the relative edge lengths shown in Figure II.*1*; and they ought to note both why the results make the target shape not match up with the original one (because the length of the top edges are different), and why the paradox is superficially plausible (because those two lengths are very close: 1½ versus $\sqrt{2}$.)

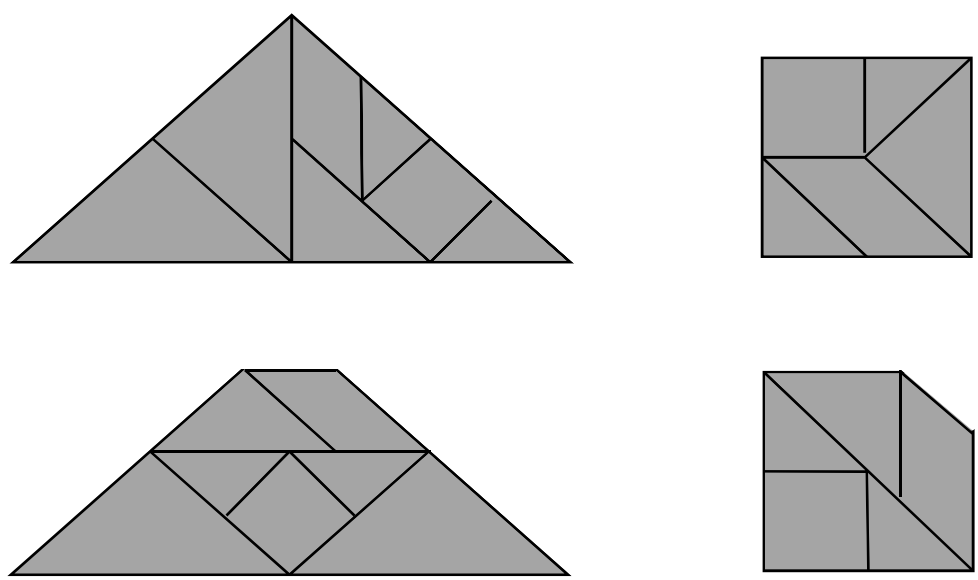


Figure V.4. Further spurious proofs using tangram rearrangement in the same basic technique

If appropriate, some very advanced wrap-up discussion might get into calculating the exact dimensions of the triangular void and the upper extrusion of the trapezoid in Figure V.*3*, or into constructing further spurious proofs based on the same fundamental trick, such as those above in Figure V.*1* from (S. Loyd, 2007, pp. 25–26).

Composing a Spurious Proof

The goal of this project is to compose an original spurious proof based on a Zeno's paradox of the Dichotomy—the proposition that motion toward a destination is impossible because the moving object must first traverse half the distance, but a quarter of the distance before or after that, and so forth ad infinitum. The advantage of using this as a starting point is that the general argument is already worked out in Zeno's philosophical exposition; composing a proof just requires translating the philosophy into mathematical language. In addition, Zeno's paradoxes are all fairly similar to each other, so if students are already familiar with the mathematical version of one paradox, then working out the mathematical version of another should be fairly familiar. Given that limits will be involved, the exercise would be suitable for the pre-calculus level or above.

The most likely context for this activity would be group exercise with the teacher leading the entire class through it, since the critical step with Zeno always relates to limits, and these tend to give students difficulty. However, it could also potentially be an individual assignment in an elective course for advanced students. Beyond the usual value of a spurious-proof exercise, this activity should have additional value as an exercise in mathematical reading and writing.

Background

Before composition of the proof begins, students should have three bits of background: Some general knowledge of Zeno and his thought, a presentation of Zeno's Achilles and the Tortoise paradox as a mathematical spurious proof, and a presentation of the Dichotomy as a purely philosophical conundrum. For general knowledge of Zeno, Heath (1921, pp. 271–83) is a reasonable choice, though his style may be rather stilted to the modern youthful ear, so a teacher may wish to look elsewhere.

A preliminary look at a mathematical version of Achilles and the Tortoise will be helpful because the two paradoxes operate on similar principles: Achilles and the Tortoise breaks the motion of two bodies down into geometrically decreasing intervals in an attempt to show that reaching a goal is impossible, while the Dichotomy breaks down the motion of just one body into geometrically decreasing intervals for the same purpose. With the similarity between the two arguments—and with the Dichotomy in fact being the simpler of the two—some prior familiarity with Achilles and the Tortoise should go a long way toward making the Dichotomy tractable. If this background is to be given out to students in written form, the exposition of Achilles and the Tortoise on page 13 might be utilized; of course, teachers might have their own preferred sources, as well. If it is desired to make the presentation of Achilles and the Tortoise interactive, the treatment at https://dysproof.blog/proofs/zenos-paradoxes/achilles-and-the-tortoise/ ("Dysproof," 2018) could be utilized. In this instance, the online material might be covered by the class together, or by the students at home in a flipped-classroom framework. In going over the background material, or checking back with the students after they have done so, the teacher should take care that students have fully understood Achilles and the Tortoise, including both fallacies involved and how they are concealed in the proof. The limit fallacy, in particular, might be difficult to explain, but there is a ready example that will be relatively fresh in the students' minds: the proposition that $.\overline{9}$ equals one. In fact, $.\overline{9}$ is not an actual decimal, but the limit of the sequence {.9, .99, .999, .9999, . . . }; so when it is said " $.\overline{9}$ equals one," what is really meant is '{.9, .99, .999, .9999, . . . } converges to one,' which is a far less fraught proposition. Of course, students generally learn decimals long before they learn limits, so the overscore notation is likely about the best that can be done under the circumstances; but students may be relieved to learn that their old teachers really were putting one over on them.

In giving students a philosophical presentation of the Dichotomy, one thing to keep in mind is that the Dichotomy is sometimes presented as dividing up the intervals already travelled, as in Heath (1921, p. 275); and sometimes as dividing up the intervals left to be travelled, as in Gardner (1982, p. 143) or Northrop (1944, pp. 119–23). The latter will make the subsequent proof writing much easier since a key step in the proof is to add up the time elapsed in travelling the subdivided intervals: if students start out looking at intervals previously travelled, they will then have to switch focus and recalculate around the intervals left to be travelled in order to avoid double-counting. The presentations focused on remaining intervals tend to have one foot in the math already, so if a teacher does not have a favorite philosophical presentation that is suited for later math work without already going in to it, a good presentation might look something like this:

‘Helen and Paris set out from Mycenae by ship, heading east toward Troy. Before they reach Troy, they must first reach the half-way point; but once they reach the half-way point, they must then travel half the remaining distance, then half the distance left after that, and their voyage can be divided up like this without end. Thus, they must sail an infinite number of spans along their way to Troy; but this will take them an infinite length of time, and they will never reach their destination.’

Composing the Proof

The proof’s composition could be done by the class all together, in smaller groups, or by students working alone as a take-home exercise. In the last case, given the difficulty students often have with limits, it would likely work best as a multiday assignment with the opportunity to ask questions before it is completed. The Dichotomy is a bit dry, and the teacher may want to dress it up with some characters, as Achilles and the Tortoise already is.

Some key points to remember as the teacher leads the class or group discussion, or answers students’ questions and assesses their work are:

Infinitely many intervals. One step in the proof must indicate that the distance to the goal is broken down into infinitely many intervals (of distance and time). This is the step that violates the limit fallacy.

Infinite time. The proof’s conclusion is that the object will never reach its destination, which means mathematically that an infinite time elapses before the destination is reached. The last step, then, should be that there is a series of time intervals that add up to an infinite sum.

Adding up the series of intervals. The simplest way to get to this final step is to argue that the distance is broken up into an infinite number of intervals, each of these takes a positive time to traverse, and so an infinite number of positive terms must add up to infinity. This, of course, is belied by the well-known geometric series formula; but this fact should be concealed by presenting the summation in a way that does not immediately call the formula to mind.

Resulting Proof

Taking all this into account, the proof produced should look something like this:

Helen and Paris leave Mycenae for Troy, a journey of positive distance D; their ship travels at positive speed s. We prove that it will take forever to reach their destination:

Before they reach Troy, they must first reach the half-way point, at a distance we define as d_1:

$$d_1 \equiv \frac{D}{2}. \tag{IV.1}$$

Having reached this point, they must then sail half the remaining distance, a span we can define as d_2:

$$d_2 \equiv \frac{d_1}{2}. \tag{IV.2}$$

As the voyage continues, so does this process of traversing half the remaining distance; and, since each such interval d_n is half the length of the last, we can calculate it in terms of D:

$$d_n = \frac{d_{n-1}}{2} = \frac{\frac{d_{n-2}}{2}}{2} = \frac{d_{n-2}}{4} = \frac{\frac{d_{n-3}}{4}}{2} = \frac{d_{n-3}}{8} = ... = \frac{D}{2^n}. \qquad \text{(IV.3)}$$

Since distance equals rate multiplied by time, knowing the ship's speed s allows us to calculate the time t_n that it takes to cover the distance d_n:

$$t_n = \frac{D_n}{s} = \frac{D}{s \cdot 2^n}. \qquad \text{(IV.4)}$$

The total time T of their journey must equal the sum of the times elapsed covering all the intermediate intervals, which go on indefinitely:

$$T = t_1 + t_2 + t_3 + ... \qquad \text{(IV.5)}$$

However, since there are an infinite number of sub-interval times $\{t_n\}$, and each is greater than zero by Equation (IV.4), the total time for their voyage must be infinite:

$$T = t_1 + t_2 + t_3 + ... \to \infty. \qquad \text{(IV.6)}$$

In this writing, the limit fallacy is incorporated in step (IV.5), where the total time is divided into an infinite number of sub-intervals; and the geometric-series fallacy is incorporated into step (IV.6), where an infinite number of terms is assumed to have an infinite sum; though, of course, student write-ups may vary somewhat from this presentation.

Chapter VI.

Conclusion

As these examples illustrate, spurious proofs can be the foundation of rather creative exercises in the mathematics classroom. This makes it particularly unfortunate that they have not generally been utilized extensively in the typical curriculum up until now.

It is doubly unfortunate in that spurious proofs seem extraordinarily well tailored to a number of goals in mathematics education with which it has otherwise proven more or less difficult to find classroom tasks to serve as building blocks toward those end goals. If we seek long-form problems to break students out of mechanistic, routinized math habits, there are many out there; but some of the best are those that surprise students to the point that their old problem-solving routine simply doesn't work anymore.

That same principle of cognitive conflict helps spurious proofs be well-keyed toward teaching the process of proof itself, as does the fact that they represent an opportunity to analyze proofs without having to compose them. Sometimes, students will only see the need for a formal proof when their old intuitive ways of grasping a situation no longer give a conclusion that makes sense. Over recent decades, math-education theorists have insisted that proof is a concept at the very core of mathematics, and that it must be taught somewhere in the math curriculum if students are to develop a full understanding of the discipline; but classroom teachers on the front lines have found relatively few opportunities to incorporate proof into their lesson plans effectively. Since spurious proofs appear to be one of the few sorts of exercises that can do so, it is regrettable that they are not utilized in this capacity more often.

Cultivating a productive disposition in math students is another area where an important goal in math education would seem to be uniquely well served by spurious proofs, but where they have largely been overlooked to date. This aspect of mathematical proficiency has been part of one of the leading analytical frameworks in the math-education world for a number of years, and yet it remains under-researched in academia and under-emphasized in the classroom. Even within the concept of productive disposition, some topics are better researched than others. The aspects of productive disposition comprising students' beliefs in their own self-efficacy as budding mathematicians are fairly well studied, especially when they intersect with demographic concerns. The aspects of productive disposition relating to students' beliefs about the cogency and sensibility of math itself receive less attention—even though a certain amount of research suggests that the mathematical aspects are indispensable, and that the mathematical and personal aspects are intertwined in surprising ways.

Productive disposition in general, and its mathematical aspects in particular, are thus underappreciated, and one might easily suspect that a major reason for this is that it is hard to find classroom activities that are specifically geared toward this goal. Spurious proofs would seem to be one such activity: with a proof's absurd conclusion, the consistency of the mathematical system is threatened, and a student can reaffirm that consistency only by actively going through the process of finding the false step in the proof. In yet another instance, a valued goal of math education has few activities that will support it, but spurious proofs are prominent among them.

In the end, then, it is disappointing that they are not put to use in the classroom in service of these objectives.

References

A bibliographic note: a number of works listed below are photoreproductions; or unabridged republications, with or without a new title. Photoreproductions are listed under their original dates and publishers, each with a parenthetical edition note identifying the photoreproduction date and publisher. Unabridged republications with original titles are listed under their original dates and new publishers, each with an edition note comprising the new publication year and the abbreviation "repb." Unabridged republications with new titles are listed as new works.

Aberdein, A. (2007a). Fallacies in mathematics. *Proceedings of the British Society for Research into Learning Mathematics*, *27*(3), 1–6.

Aberdein, A. (2007b). The informal logic of mathematical proof. In B. Van Kerkhove & J. P. Van Bendegem (Eds.), *Perspectives on mathematical practices: Bringing together philosophy of mathematics, sociology of mathematics, and mathematics education* (pp. 135–51). Dordrecht: Springer-Verlag. http://doi.org/10.1007/0-387-29831-2_4

Aberdein, A. (2009). Mathematics and argumentation. *Foundations of Science*, *14*(1–2), 1–8. http://doi.org/10.1007/s10699-008-9158-3

Anderson, L. (2007, September 21). Comment on John Edwards and—Daniel Biss? *Chicago Reader*. Chicago.

Atiyah, M., Borel, A., Chaitin, G. J., Friedan, D., Glimm, J., Gray, J. J., … Zeeman, E. C. (1994). Responses to "Theoretical mathematics: toward a cultural synthesis of mathematics and theoretical physics", by A. Jaffe and F. Quinn. *Bulletin of the American Mathematical Society*, *30*(2), 178–207.

Balacheff, N. (2010). Bridging Knowing and Proving in Mathematics: A Didactical Perspective. In G. Hanna, H. N. Jahnke, & H. Pulte (Eds.), *Explanation and Proof in Mathematics: Philosophical and Educational Perspectives* (pp. 115–35). New York: Springer.

Ball, D. L., Hoyles, C., Jahnke, H. N., & Movshovitz-Hadar, N. (2002). The teaching of proof. *Arxiv Preprint math0305021*, *III*(section 2), 1–12. Retrieved from http://eprints.ioe.ac.uk/1801/

Barbeau, E. J. (1983). Rigorous Proof in Mathematics Education. *Interchange*, *14*(4), 163–67.

Barbeau, E. J. (1990). 3 Faces of Proof. *Interchange*, *21*(1), 24–27.

Barbeau, E. J. (2000). *Mathematical fallacies, flaws, and flimflam.* Washington, DC: Mathematical Association of America.

Barbeau, E. J. (2013). *More Fallacies, Flaws & Flimflam*. Mathematical Association of America.

Barker, K., Burdick, D., Stek, J., Wessel, W., Youngblood, R., Hindson, E., & Mitchell, D. (Eds.). (2002). *KJV Study Bible*. Grand Rapids, MI: Zondervan.

Bassarear, T. (1990). An Examination of the Influence of Attitudes and Beliefs on Achievement in a College Developmental Mathematics Course. *Research & Teaching in Developmental Education*, *7*(2), 43–56.

BBC News. (2018). "Unsolvable" exam question leaves Chinese students flummoxed. Retrieved from http://www.bbc.com/news/world-asia-china-42857864

Bell, E. T. (1937). *Men of Mathematics* (Kindle). New York: Simon & Schuster.

Blazar, D., & Kraft, M. A. (2016). Teacher and Teaching Effects on Students Attitudes and Behaviors. *Educational Evaluation and Policy Analysis*, *39*(1), 146–70. http://doi.org/10.3102/0162373716670260

Boaler, J. (1998). Open and Closed Mathematics: Student Experiences and Understandings. *Journal for Research in Mathematics Education*, *29*(1), 41–62.

Boero, P. (2007). *Theorems in school: from history, epistemology and cognition to classroom practice*. Rotterdam: Sense.

Bradis, V. M., Minkovskii, V. L., & Kharcheva, A. K. (1963). *Lapses in Mathematical Reasoning*. (E. A. Maxwell, Ed., J. J. Schorr-Kon, Trans.) (1999 repb.). Mineola, NY: Dover.

Bunch, B. (1997). *Mathematical Fallacies and Paradoxes*. Mineola, NY: Dover.

De Morgan, A. (1915). *A Budget of Paradoxes*. (D. E. Smith, Ed.) (2d ed.). Chicago: Open Court.

de Villiers, M. (2010). Experimentation and Proof in Mathematics. In G. Hanna, H. N. Jahnke, & H. Pulte (Eds.), *Explanation and Proof in Mathematics: Philosophical and Educational Perspectives* (pp. 205–22). New York: Springer.

Dee, J. (2007, September 13). Comment on John Edwards and—Daniel Biss? *Chicago Reader*. Chicago.

Dewey, J. (1933). *How We Think: A Restatement of the Relation of Thinking to the Educative Process*. Boston: D.C. Heath and Co.

Dubnov, Y. S. (1963). Mistakes in Geometric Proofs. In T. M. Switz & L. Lange

(Trans.), *Proof in Geometry with "Mistakes in Geometric Proofs"* (2006 repb.). Mineola, NY: Dover.

Dudeney, H. E. (1908). *The Canterbury Puzzles, and Other Curious Problems* (2017 Scholar Select phrp.). New York: E.P. Dutton & Co.

Dudeney, H. E. (1917). *Amusements in Mathematics* (2017 repb.). Middletown, DE: Palmera.

Dudeney, H. E. (1967). *536 Curious Problems and Puzzles* (1995 repb.). New York: Barnes & Noble.

Dweck, C. S. (1986). Motivational Processes Affecting Learning. *American Psychologist*, *41*(10), 1040–48. http://doi.org/10.1037/0003-066X.41.10.1040

Dysproof. (2018). Retrieved from dysproof.blog

Ellis, A. B., Bieda, K., & Knuth, E. (2012). *Developing Essential Understanding of Proof and Proving: Grades 9–12*. Reston, VA: National Council of Teachers of Mathematics.

Eveleth, R. (2013, July). If Women Assume Fake Names, They Do Better on Math Tests. *Smithsonian*. Retrieved from https://www.smithsonianmag.com/smart-news/if-women-assume-fake-names-they-do-better-on-math-tests-6390944/

Farlow, S. J. (2014). *Paradoxes in Mathematics*. Mineola, NY: Dover.

Fetisov, A. I. (1963). Proof in Geometry. In T. M. Switz & L. Lange (Trans.), *Proof in Geometry with "Mistakes in Geometric Proofs"* (2006th ed.). Mineola, NY: Dover.

Fitzgerald, J. F. (1996). Proof in Mathematics Education. *Journal of Education*, *178*(1), 35–45.

Fixx, J. (1972). *Games for the Super-intelligent*. Garden City, NY: Doubleday.

Frank, M. L. (1988). Problem solving and mathematical beliefs. *The Arithmetic Teacher*, *35*(5), 32–34. Retrieved from http://kmi.open.ac.uk/publications/techreport/kmi-10-03

Gardner, M. (1956). *Mathematics, Magic and Mystery* (Kindle). New York: Dover.

Gardner, M. (1978). *aha! Insight* (2006 repb.). Washington, DC: Mathematical Association of America.

Gardner, M. (1982). *aha! Gotcha* (2006 repb.). Washington, DC: Mathematical Association of America.

Gardner, M. (1983). *Wheels, Life and Other Mathematical Amusements*. New York:

W.H. Freeman & Co.

Gardner, M. (1987). *The Second Scientific American Book of Mathematical Puzzles and Diversions*. Chicago: University of Chicago Press.

Gardner, M. (1988a). *Hexaflexagons and Other Mathematical Diversions: The First Scientific American Book of Mathematical Puzzles and Games*. Chicago: University of Chicago Press.

Gardner, M. (1988b). *Time Travel and Other Mathematical Bewilderments*. New York: W.H. Freeman & Co.

Garofalo, J. (1989). Beliefs and their Influence on Mathematics Performance. *Mathematics Teacher*, *82*(7), 502–05. Retrieved from http://www.jstor.org/stable/27966379

Gatward, R. (2011). Is there a need for proof in secondary mathematics education? *Mathematics Teaching*, *225*(November), 6–8.

Goodrich, A. T. (1817). *The New and Fashionable Chinese Puzzle*. New York: A.T. Goodrich. Retrieved from http://nrs.harvard.edu/urn-3:hul.ebookbatch.EAISS_batch:aas04303056

Goodrich, A. T. (1818). *The Fashionable Chinese Puzzle*. New York: A.T. Goodrich. Retrieved from http://nrs.harvard.edu/urn-3:hul.ebookbatch.EAISS_batch:aas04297042

Grady, M. (2016). Whatever Happened to Productive Disposition? *Mathematics Teaching in the Middle School*, *21*(9), 516–18.

Hadas, N., Hershkowitz, R., & Schwarz, B. B. (2000). The role of contradiction and uncertainty in promoting the need to prove in dynamic geometry environments. *Educational Studies in Mathematics*, *44*, 127–50. http://doi.org/10.1023/A:1012781005718

Hanna, G. (1989). More than Formal Proof. *For the Learning of Mathematics*, *9*(1), 20–23.

Hanna, G. (1990). Some pedagogical aspects of proof. *Interchange*, *21*(1), 6–13.

Hanna, G. (1995). Challenges to the Importance of Proof. *For the Learning of Mathematics*, *15*(3), 42–49.

Hanna, G. (2000a). A Critical Examination of Three Factors in the Decline of Proof. *Interchange*, *31*(1), 21–33.

Hanna, G. (2000b). Proof, Explanation and Exploration: An Overview. *Educational Studies in Mathematics*, *44*(1/2), 5–23.

http://doi.org/10.1023/A:1012737223465

Hanna, G., & Barbeau, E. (2008). Proofs as bearers of mathematical knowledge. *ZDM Mathematics Education*, *40*, 345–53.

Hanna, G., & de Villiers, M. (2008). ICMI Study 19: Proof and proving in mathematics education. *ZDM Mathematics Education*, *40*, 329–36.

Hanna, G., & de Villiers, M. (Eds.). (2012). *Proof and Proving in Mathematics Education: The 19th ICMI Study*. New York: Springer.

Hanna, G., & Jahnke, H. N. (1993). Proof and Application. *Educational Studies in Mathematics*, *24*(4), 421–38. http://doi.org/10.1007/BF01273374

Hanna, G., & Jahnke, H. N. (1996). Proof and proving. In A. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education* (pp. 877–908). Dordrecht: Kluwer.

Hanna, G., Jahnke, H. N., & Pulte, H. (Eds.). (2010). *Explanation and Proof in Mathematics: Philosophical and Educational Perspectives*. New York: Springer-Verlag.

Hanna, G., & Sidoli, N. (2007). Visualisation and proof: A brief survey of philosophical perspectives. *ZDM - International Journal on Mathematics Education*, *39*(1–2), 73–78. http://doi.org/10.1007/s11858-006-0005-0

Hartswick, F. G. (1929). Ride with Youth. *Journal of Education*, *109*(18), 508–09.

Heath, T. L. (1921). *A History of Greek Mathematics* (Vol. 1). New York: Cambridge University Press.

Hemmi, K. (2006). *Approaching Proof in a Community of Mathematical Practice*. Stockholm University.

Hemmi, K. (2010). Three styles characterising mathematicians' pedagogical perspectives on proof. *Educational Studies in Mathematics*, *75*(3), 271–91. http://doi.org/10.1007/s10649-010-9256-3

Henderson, H. (2007, August 9). John Edwards and—Daniel Biss? *Chicago Reader*. Chicago.

Hocker, T. (2017). *A Study of Perceptions of Math Mindset, Math Anxiety, and View of Math by Young Adults*. Southeastern University.

Horgan, J. (1993). The Death of Proof. *Scientific American*, *269*(4), 93–103.

Huck, S. W., & Sandler, H. M. (1984a). *Statistical Illusions: Problems*. New York: Harper & Row.

Huck, S. W., & Sandler, H. M. (1984b). *Statistical Illusions: Solutions*. New York: Harper & Row.

Jacobson, E. D. (2017). Field Experience and Prospective Teachers' Mathematical Knowledge and Beliefs. *Journal for Research in Mathematics Education*, *48*(2), 148–90. Retrieved from http://www.jstor.org/stable/10.5951/jresematheduc.48.2.0148%5Cnhttp://www.jstor.org/stable/10.5951%5Cnhttp://about.jstor.org/terms

Jacobson, E., & Kilpatrick, J. (2015). Understanding teacher affect, knowledge, and instruction over time: an agenda for research on productive disposition for teaching mathematics. *Journal of Mathematics Teacher Education*, *18*(5), 401–06. http://doi.org/10.1007/s10857-015-9316-9

Jaffe, A., & Quinn, F. (1993). "Theoretical mathematics'': Toward a cultural synthesis of mathematics and theoretical physics. *Bulletin of the American Mathematical Society*, *29*(1), 1–13. http://doi.org/10.1090/S0273-0979-1993-00413-0

Jaffe, A., & Quinn, F. (1994). Response to comments on "Theoretical mathematics'', *30*(2), 4. http://doi.org/10.1090/S0273-0979-1994-00506-3

Jones, J. W. (2016). *A Quantitative Study: The Relationship between School-Wide Instructional Practices, Teacher Beliefs, and Growth Mindset and Value-Added Student Growth in Elementary Mathematics for Grades 3–5 of Pennsylvania December 2016*. Indiana University of Pennsylvania.

Jong, C., & Hodges, T. E. (2015). Assessing attitudes toward mathematics across teacher education contexts. *Journal of Mathematics Teacher Education*, *18*(5), 407–25. http://doi.org/10.1007/s10857-015-9319-6

Klymchuk, S., Kachapova, F., Schott, D., Sauerbier, G., & Kachapov, I. (2011). Using paradoxes and counterexamples in teaching probability: a parallel study. *Mathematics Teaching Research Journal Online*, *5*(1), 50–70.

Klymchuk, S., & Staples, S. (2013). *Paradoxes and Sophisms in Calculus*. Washington: Mathematical Association of America.

Kondratieva, M. (2009a). Geometrical Sophisms and Understanding of Mathematical Proofs. In F.-L. Lin, F.-J. Hsieh, G. Hanna, & M. de Villiers (Eds.), *Proceedings of the ICMI Study 19 Conference: Proof and Proving in Mathematics Education* (Vol. 2, pp. 3–8). Taipei: Department of Mathematics, National Taiwan Normal University.

Kondratieva, M. (2009b). Understanding Mathematics through Resolution of Paradoxes. In B. Sriraman, V. Freiman, & N. Lirette-Pitre (Eds.), *Interdisciplinarity, Creativity, and Learning* (pp. 61–74). Charlotte, NC: Information Age.

Larvor, B. (2010). Authoritarian Versus Authoritative Teaching: Polya and Lakatos. In G. Hanna, H. N. Jahnke, & H. Pulte (Eds.), *Explanation and Proof in Mathematics: Philosophical and Educational Perspectives* (pp. 71–83). New York: Springer-Verlag.

Lewis, J. M., Fischman, D., & Riggs, M. (2015). Defining, developing, and measuring "Proclivities for Teaching Mathematics." *Journal of Mathematics Teacher Education*, *18*(5), 447–65. http://doi.org/10.1007/s10857-015-9321-z

Lin, F.-L., Hsieh, F.-J., Hanna, G., & de Villiers, M. (Eds.). (2009a). *Proceedings of the ICMI Study 19 conference: Proof and Proving in Mathematics Education* (Vol. 2). Taipei: Department of Mathematics, National Taiwan Normal University. Retrieved from http://140.122.140.1/~icmi19/files/Volume_2.pdf

Lin, F.-L., Hsieh, F.-J., Hanna, G., & de Villiers, M. (Eds.). (2009b). *Proceedings of the ICMI Study 19 Conference: Proof and Proving in Mathematics Education* (Vol. 1). Taipei: Department of Mathematics, National Taiwan Normal University. Retrieved from http://140.122.140.1/~icmi19/files/Volume_1.pdf

Lin, F.-L., Yang, K.-L., Lo, J.-J., Tsamir, P., Tirosh, D., & Stylianides, G. (2012). Teachers' Professional Learning of Teaching Proof and Proving. In G. Hanna & M. de Villiers (Eds.), *Proof and Proving in Mathematics Education: The 19th ICMI Study* (pp. 327–46). New York: Springer.

Loyd, S. (1903). *The Eighth Book of Tan*. (P. Van Note, Ed.) (1968 repb.). New York: Dover.

Loyd, S. (1914). *Sam Loyd's Cyclopedia of 5000 Puzzles Tricks and Conundrums with Answers*. (W. "Sam" Loyd & S. Sloan, Eds.) (2007 Ishi phrp.). New York: Lamb Publishing Company.

Loyd, S. (1959). *Mathematical Puzzles of Sam Loyd*. (M. Gardner, Ed.). New York: Dover.

Loyd, S. (1960). *More Mathematical Puzzles of Sam Loyd*. (M. Gardner, Ed.). New York: Dover.

Loyd, S. (2007). *Sam Loyd's Book of Tangrams*. (P. Van Note, Ed.). Mineola, NY: Dover.

Lutzer, C. V. (2005). The Role of Proof in Mathematics and Mathematics Education. *Primus*, *15*(4), 374–84. http://doi.org/10.1080/10511970508984130

Lyamin, A. A. (1903). *Son i snovideniya. Gipnotizm. Spiritizm. Telepatiya. Yasnovideniya: Ocherk* (2012 Knizhnyy Renessans phrp.). Moscow: D.N. Kornatovsky Russian Artistic Press.

Lyamin, A. A. (1911). *Matematicheskiye paradoksy i interesnyye zadachi dlya lyubiteley matematiki* (2012 Knizhnyy Renessans phrp.). Moscow: G. Lissner and D. Sobko.

Mac Lane, S. (1997). Despite Physicists, Proof Is Essential in Mathematics. *Synthese*, *111*, 147–54. http://doi.org/10.1093/philmat/nkm010

Mamolo, A., & Zazkis, R. (2008). Paradoxes as a window to infinity. *Research in Mathematics Education*, *10*(2), 167–82. http://doi.org/10.1080/14794800802233696

Mariotti, A. (2006). Proof and proving in mathematics education. In A. Gutierrez & P. Boero (Eds.), *Handbook of research on the psychology of mathematics education: Past, present and future* (pp. 173–204). Rotterdam: Sense.

Markel, W. D. (1994). The Role of Proof in Mathematics Education. *School Science and Mathematics*, *94*(6), 291–95.

Mathematics Teaching. (1996). *Mathematics Teaching* (Vol. 155).

Maxwell, E. A. (1959). *Fallacies in mathematics*. Cambridge: Cambridge University Press.

McInroy, S. (1994). *Puzzles, Pizza, Paradoxes, Phibonacci and Fallacies*. Corning, NY: Corning Community College.

Mnev, N. (2007). On D.K. Biss' papers "The homotopy type of the matroid Grassmannian" and "Oriented matroids, complex manifolds, and a combinatorial model for BU," *158*(2003), 2. Retrieved from http://arxiv.org/abs/0709.1291

Movshovitz-Hadar, N. (1988). Stimulating presentation of theorems followed by responsive proofs. *For the Learning of Mathematics*, *8*(2), 12–19.

Movshovitz-Hadar, N., & Hadass, R. (1990). Preservice education of math teachers using paradoxes, *21*(3), 265–87.

Movshovitz-Hadar, N., & Hadass, R. (1991). More about mathematical paradoxes in preservice teacher education. *Teaching and Teacher Education*, *7*(1), 79–92. http://doi.org/10.1016/0742-051X(91)90059-X

Movshovitz-Hadar, N., & Webb, J. (1998). *One Equals Zero and Other Mathematical Surprises*. Emeryville, CA.

National Research Council. (2001). *Adding It Up: Helping Children Learn Mathematics*. (J. Kilpatrick, J. Swafford, & B. Findell, Eds.). Washington, DC: National Academy Press.

Noormohamed, F. (2010). Prove It! Importance of Proof Thinking in Mathematical Education. *Education Canada*, *46*(1), 25–26.

Northrop, E. P. (1944). *Riddles in Mathematics* (2014 repb.). Mineola, NY: Dover.

Nufer, D. (2017). *The Tangram Fury Paradox Book*. Middletown, DE: Peepsock Press.

Paul, A. M. (2012, October 7). It's Not Me, It's You. *New York Times*. New York. Retrieved from http://www.nytimes.com/2012/10/07/opinion/sunday/intelligence-and-the-stereotype-threat.html

Philipp, R. A., & Siegfried, J. M. (2015). Studying productive disposition: the early development of a construct. *Journal of Mathematics Teacher Education*, *18*(5), 489–99. http://doi.org/10.1007/s10857-015-9317-8

Philippou, G. N., & Christou, C. (1998). The Effects of a Preparatory Mathematics Program in Changing Prospective Teachers' Attitudes Towards Mathematics. *Educational Studies in Mathematics*, *35*(2), 189–206.

Pohlers, W. (1994). *Proof Theory: An Introduction*. (A. Dold, B. Eckmann, & F. Takens, Eds.) (2d prntng). New York: Springer-Verlag.

Polya, G. (1954a). Induction and Analogy in Mathematics. In *Mathematics and Plausible Reasoning* (2014 repb., Vol. 1). Mansfield Centre, CT: Martino.

Polya, G. (1954b). Patterns of Plausible Inference. In *Mathematics and Plausible Reasoning* (2014 repb., Vol. 2). Mansfield Centre, CT: Martino.

Rav, Y. (1999). Why do we prove theorems? *Philosophia Mathematica*, *7*(1), 5–41. http://doi.org/10.1093/philmat/7.1.5

Read, R. C. (1965). *Tangrams: 330 Puzzles*. New York: Dover.

Reid, D. (1996). Teaching and the role of proof. *Mathematics Teaching*, *155*(June), 23–26.

Reid, D. A., & Knipping, C. (2010). *Proof in mathematics education: research, learning, and teaching.* Rotterdam: Sense.

Rezaian, J. (2018, January 31). This Chinese math problem has no answer. Actually, it has a lot of them. Retrieved from https://www.washingtonpost.com/news/worldviews/wp/2018/01/31/this-chinese-math-problem-has-no-answer-actually-it-has-a-lot-of-them/?utm_term=.7541b95a0cde

Sarcone, G. (2017). Tangramagic Puzzle. Retrieved from https://www.giannisarcone.com/wp/blog/2017/10/02/tangramagic-puzzle/#more-164

Schoenfeld, A. H. (Ed.). (2007). *Assessing Mathematical Proficiency*. New York: Cambridge.

Schram, P., Wilcox, S. K., Lappan, G., & Lanier, P. (1988). Changing preservice teachers' beliefs about mathematics education. In C. A. Mahers, G. A. Goldin, & R. B. Davis (Eds.), *Proceedings of the 11th conference of the North American group for the psychology of mathematics education* (Vol. 1, pp. 296–302). New Brunswick, NJ: Rutgers University.

Schutte, K. (2007). *Proof Theory*. New York: Springer-Verlag. http://doi.org/10.1007/978-3-642-25847-3

Sfondeles, T. (2017, September 29). Are gov hopeful Biss's claims of math prowess pi in the sky? *Chicago Sun-Times*. Chicago. Retrieved from https://chicago.suntimes.com/news/are-gov-hopeful-bisss-claims-of-math-prowess-pi-in-the-sky/

Siegfried, J. M. (2012). *The Hidden Strand of Mathematical Proficiency: Defining and Assessing for Productive Disposition in Elementary School Teachers' Mathematical Content Knowledge*. UC San Diego.

Slocum, J. (2003). *The Tangram Book*. New York: Sterling.

Slocum, J. (2006). *Sam Loyd's Most Successful Hoax*. Retrieved from https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=9&ved=0ahUKEwj19vrsqorYAhUIerwKHcENDXIQFghDMAg&url=http%3A%2F%2Fwww.indiana.edu%2F~liblilly%2Fcollections%2Foverview%2Fpuzzle_docs%2FSam_Loyd_Successful_Hoax.pdf&usg=AOvVaw3Ea1NXR2ju4zwNvbfsIQg

Slocum, J., & Sonneveld, D. (2006). *The 15 Puzzle Book*. Beverly Hills, CA: Slocum Puzzle Foundation.

Smullyan, R. M. (1978). *What Is the Name of This Book?* (2011 repb.). Mineola, NY: Dover.

Sriraman, B. (2009). Mathematical Paradoxes as Pathways into Beliefs and Polymathy. In B. Sriraman, V. Freiman, & N. Lirette-Pitre (Eds.), *Interdisciplinarity, Creativity, and Learning* (pp. 75–93). Charlotte, NC: Information Age.

Stodolsky, S. S. (1985). Telling Math: Origins of Math Aversion and Anxiety. *Educational Psychologist*, *20*(3), 125–33. http://doi.org/10.1207/s15326985ep2003

Stylianides, A. J., & Harel, G. (Eds.). (2018a). *Advances in Mathematics Education*

Research on Proof and Proving: An International Perspective. New York: Springer.

Stylianides, A. J., & Harel, G. (2018b). Preface. In A. J. Stylianides & G. Harel (Eds.), *Advances in Mathematics Education Research on Proof and Proving: An International Perspective* (pp. v–viii). New York: Springer.

Stylianides, A. J., & Stylianides, G. J. (2014). Impacting positively on students' mathematical problem solving beliefs: An instructional intervention of short duration. *Journal of Mathematical Behavior*, *33*(1), 8–29. http://doi.org/10.1016/j.jmathb.2013.08.005

Stylianides, G. J., & Stylianides, A. J. (2009). Facilitating the Transition from Empirical Arguments to Proof. *Journal for Research in Mathematics Education*, *40*(3), 314–52. http://doi.org/10.2307/40539339

Stylianou, D. A., Blanton, M. L., & Knuth, E. J. (Eds.). (2009). *Teaching and learning proof across the grades: a K–16 perspective*. New York: Routledge.

Swars, S. L., Smith, S. Z., Smith, M. E., & Hart, L. C. (2009). A longitudinal study of effects of a developmental teacher preparation program on elementary prospective teachers' mathematics beliefs. *Journal of Mathematics Teacher Education*, *12*(1), 47–66.

Szpiro, G. G. (2010). Brilliant but Fallible. In E. Burke (Trans.), *A Mathematical Medley: Fifty Easy Pieces on Mathematics* (pp. 97–99). Providence, RI: The American Mathematical Society.

Takeuti, G. (1975). *Proof Theory*. New York: Elsevier.

Tall, D. (Ed.). (1991). *Advanced Mathematical Thinking*. Boston: Kluwer.

Tall, D., Yevdokimov, O., Koichu, B., Whiteley, W., Kondratieva, M., & Cheng, Y.-H. (2012). Cognitive Development of Proof. In G. Hanna & M. de Villiers (Eds.), *Proof and Proving in Mathematics Education: The 19th ICMI Study* (pp. 13–49). New York: Springer.

Thomas, R. (1998). An Update on the Four-Color Theorem. *Notices of the AMS*, *45*(7), 848–59. http://doi.org/10.1097/00019052-200112000-00018

Thurston, W. P. (1994). On Proof and Progress in Mathematics. *Bulletin of the American Mathematical Society*, *30*(2), 161–77.

Troelstra, A. S., & Schwichtenberg, H. (2000). *Basic Proof Theory* (2d ed.). New York: Cambridge University Press.

Van Note, P. (1973). *Tangrams: Picture-Making Puzzle Game* (2nd ed.). Rutland, VT:

Tuttle Press.

Verschaffel, L., Greer, B., & De Corte, E. (2000). *Making Sense of Word Problems*. Exton, PA: Swets & Zeitlinger.

Voronets, A. M. (1930). *Vil'noï hodyny : zbirka rebusiv, sharad, lohohryfiv, matematychnykh zavdan', tanhram, stomakhion ta inshykh rozvah*. Kharkiv: Knyhospilka.

Web of Science. (2018). Retrieved from http://apps.webofknowledge.com.ezp-prod1.hul.harvard.edu/WOS_CitedReferenceSearch_input.do?SID=7A1t4sE2TpqDdA3hNkS&product=WOS&search_mode=CitedReferenceSearch

Wessen, K. (2015). Bamboozlers : Encouraging Mathematical Thinking Through Puzzles and Paradox. In *MANSW Annual Conference*. Retrieved from http://thewessens.net/ClassroomApps/Bamboozlers (MANSW 2015).pdf

Zaslavsky, O. (2005). Seizing the opportunity to create uncertainty in learning mathematics. *Educational Studies in Mathematics*, *60*(3), 297–321. http://doi.org/10.1007/s10649-005-0606-5

Zaslavsky, O., Nickerson, S. D., Stylianides, A. J., Kidron, I., & Winicki-Landman, G. (2012). The Need for Proof and Proving: Mathematical and Pedagogical Perspectives. In G. Hanna & M. de Villiers (Eds.), *Proof and Proving in Mathematics Education: The 19th ICMI Study* (pp. 215–29). New York: Springer.