

The Authority of the Incomprehensible

B. Mazur

February 16, 2014

You may not know what *Abracadabra* means, but you very well feel its magical force, admit it! Its effect only gains from the obscurity of the incantation. It is true, of course, that *ipsa scientia potestas est*, (Latin—by the way—for “knowledge itself has powers”) but being confronted with something that purports to be wisdom and is Greek to you, is even more powerful.

In a word, we humans are prone to take certain *incomprehensible* assertions as carrying some kind of evidentiary authority *because of* their incomprehensibility, and irrespective of the content of whatever messages those assertions were meant to communicate.

Children are constantly showered with words, phrases, usages—that mean both nothing and everything to them—from those beings that tower over them. These are utterances that have presumably potent meaning to the adult world, and—deliciously—can be repeated, in a kind of Bayesian language game, to see what power one derives from proclaiming them.

I remember that the first time in my life I heard the word *Chinatown* mentioned blithely by some adult, I knew nothing of its referent, and was in awe that a presumably visitable town(?) could miniaturize, distill, and encompass a vast country and language with an inaccessible script: I felt compelled to use that word, *Chinatown*, repeatedly, in whatever context arose, no matter what it meant, for a full week after I heard it, to possess the sheer power of it.

And what child doesn’t make constant magic by singing that—incomprehensible—single-word song: “Why?”?

Even the scrambling of sense into nonsense is a joy:

Mairzy doats and dozy doats
and liddle lamzy divey.
A kiddley divey too,
wooden shoe?

I’m a mathematician. I love math as a sterling example of how far and how deep you can get just by *organized thinking*. I love its beauty, its profundity, its surprises. I love that it finds ways

to help us with every aspect of our existence: mathematics lives in the very shape of music, the construction of cathedrals, the conception of internet. Mathematics expresses our general lust for understanding.

So it makes me cringe when I see the strange appearance of a piece of mathematics in some argument, or application, where the only function it plays is to be *not really understood* and—thereby!— to convey a level of gravitas that the argument ‘in clear’ wouldn’t have. Sometimes a tidbit of math is paraded rather than used. I suppose that math is quite a dependable obscurifier, weighty enough to silence any objections, even if it is put on the table as a never-to-be-cashed-in rhetorical chip¹.

My father-in-law was a tax lawyer who loved to read articles about tax in professional journals, but would be utterly stymied when a piece of math appeared. These roadblocks to his reading occurred quite often, and the two of us made a habit of collecting examples where the math said absolutely nothing other than what the English sentence preceding it in the text said perfectly well. (Usually the English would refer to some kind of exponential growth and the math would be a trumpeted differential equation with none of the variables or parameters defined.)

These examples made quite curious reading to me, a mathematician, as if the author, wanting to be helpful to **me**, suddenly worried that since math was (presumably) my native language, my English-comprehension needed to be given a boost with an infusion of math-talk from time to time.

It would be—reversing roles—as if I—thinking all of a sudden that I might have some Spanish readers—would follow this current sentence in this text with a perfect translation of it into Spanish, but have the gall to proclaim the translation as an extra argument in favor of my conclusions.

We model the world in order to understand it. This is wonderful. And there is no model that isn’t driven by mathematics, or at the very least is shaped by a mathematical sensibility.

But there are dangers. Some mathematical models have been blindly used—their presuppositions as little understood as any legal fine print one “agrees to” but never reads—with faith in their trustworthiness. The very arcane nature of some of the formulations of these models might have contributed to their being given so much credence. If so, we mathematicians have an important mission to perform: to help people who wish to think through the fundamental assumptions underlying models that are couched in mathematical language, making these models intelligible, rather than (merely) formidable Delphic oracles.

There is currently a vigorous public discussion about models, their role in Economics and the issue of whether or not their utility should be measured merely on their ability to predict outcomes; and whether they *do* have the ability to predict, all that well. Discussion about the roles and limitations, as well as the essential virtues of models is a good thing. See, for example, *Sunday Dialogue: Economics’ Ups and Downs* published in the New York Times (August 31, 2013). In

¹On the other hand, as Yuri Tschinkel pointed out to me, the journal *Science* (July 20, 2012) surveyed 649 papers in ecology and evolution in 1998, noting that each mathematical equation in the main text of the paper was associated with a 28% decrease of the citation rate. This seems to point to some damping mechanism in the mere use of math-for-rhetoric.

particular, see the letter there by Eric Maskin who argues for economic models to play the role, effectively, of thought-laboratories aimed either at explanation even when prediction is impossible. In other writings, Maskin discusses economic models as *parables*. These are useful attitudes, I believe, and should be more commonly understood. Related to this, consider this blog-note of Paul Krugman (New York Times, August 21, 2013):

Noah Smith (<http://noahpinionblog.blogspot.com/2013/08/a-few-words-about-math.html>) has a fairly caustic meditation on the role of math in economics, in which he says that its nothing like the role of math in physics and suggests that it's mainly about doing hard stuff to prove that you're smart.

I share much of his cynicism about the profession, but I think he's missing the main way (in my experience) that mathematical models are useful in economics: used properly, they help you think clearly, in a way that unaided words can't.

Take the centerpiece of my early career, the work on increasing returns and trade. The models I and others used were, in a way, typical of economics: clearly untrue assumptions (symmetric constant elasticity of substitution preferences; symmetric costs across products!), and involved a fair bit of work to arrive at what sounds in retrospect like a fairly obvious point: even similar countries will end up specializing in different products, and because there are increasing returns in many sectors, this will produce gains from specialization and trade. But this point was only obvious in retrospect. People in trade were not saying anything like this until the New Trade Theory models came along and clarified our thinking and language. Trust me, I was there, and went through a number of seminar experiences in which I had to bring an uncomprehending audience through until they saw the light.

The same is true for the liquidity trap. The basics of what happens at the zero lower bound aren't complicated, but people who haven't worked through small mathematical models of both the IS-LM and New Keynesian type generally get all tied up in verbal and conceptual knots.

What is true is that all too many economists have lost sight of this purpose; they treat their models as The Truth, and/or judge each others work by how hard the math is. It sounds as if Smith was taught macro by people like that. And there are a lot of people in macro, some of them fairly prominent, who are what my old teacher Rudi Dornbusch used to call "fearful plumbers" people who can push equations around, but have no sense of what they mean, and as a result say quite remarkably stupid things when confronted with real-world economic issues.

But math is good, used right.

I agree, with the emphasis on *used right!* I'm a convinced user of mathematical models of all sorts. We can't and we shouldn't avoid dealing with mathematical vocabulary in the construction of our models. But it would be sad if people's trust in the predictions of a mathematical model rested merely on its obscurity, or difficulty, or abstractness, or lack of transparency.

Trust is natural. We depend upon experts, and we trust the cloak of experts, as—given the alternatives—we probably should. And there is a fundamental finality to an actual *number* derived by experts in a manner that is beyond your judging; it represents the *bottom line* of the argument,

since what kind of discussion could come after it? Your LDL number is 150! *Waddyagonnadoaboutit?* The answer, clearly, is not to do a critique of the *risk-assessment* software that projects some statistics at you: e.g., the model’s calculation of the percentage of the population of roughly your age and with LDL=150 who will encounter heart problems in the next 10 years. You first, I guess, would just go about doing something about your cholesterol.

Of course, there are some outputs of models that have the appealing property that their procedure is devilishly difficult to justify—and whose inner workings are not easy to understand—but once they come up with an answer, the correctness of the answer is miraculously easy to check. In such a case, it doesn’t matter how flimsy the model, or how obscure the reasoning is to you that led to the discovery of the answer—it could be by ouija board! But if the answer once found is easily checked, then just check it!

An example of this—interior to mathematics—is the primordial issue of factoring whole numbers: give a modern-day computer a large number, N , having—say—300 digits, and (even if you know in advance that it is the product of two primes of roughly 150 digits each) there’s no guarantee that the computer can ‘factor’ N , that is, find its two prime factors, $N = P \cdot Q$. But suppose some supercomputer does, laboriously, manage to make the factorization and presents to you the factors P and Q . It is then very easy for any computer (super or not) to multiply them together to check that, in fact, $P \cdot Q$ equals N . Once armed with P and Q you don’t worry about the opaqueness of the procedure that generated them; you just multiply them. This huge discrepancy in speed of performing an operation versus the speed of performing the inverse of the operation is the basis of much modern cryptography².

This latter example is not the type of model-receptivity I’m focusing on in this essay. That is, these types of models may or may not be obscure, but the estimate of veracity of their output doesn’t, somehow, depend upon that obscurity.

As in the law journals I pored over with my father-in-law, the single most touted mathematical model brought as witness in various quintessentially non-mathematical studies is the *Malthusian Growth Model*. Now the Malthusian idea is extremely important in itself, as well as for historical reasons. Its very structure teaches us a lot; it is a simple starting place, to consider what mathematical models can do for you and what they can’t. And it has the advantage of providing models that are *comprehensible*, whether or not they appropriately describe the reality they are constructed to model.

²I recently witnessed such a discrepancy when I needed to factor this 204-digit number:

345269032939215803146410928173696740406844 ~
~ 815684239672101299206421451944591925694154 ~
~ 456527606766236010874972724155570842527652 ~
~ 72786877636295951962087273561220060103650 ~
~ 6871681124610986596878180738901486527

As a team of experts worked on it for over a six month period, I was privy to their frequent emails regarding progress in this project, but could understand not a single word of those emails, until the final message which presented the two factors. Once those factors were given, any reasonable computer can simply multiply them together to check that their product is the displayed 204-digit number.

Thomas Malthus (*An Essay on the Principle of Population* 1798) gives himself two starting postulates:

First, that food is necessary to the existence of man.

Secondly, that the passion between the sexes is necessary and will remain nearly in its present state.

What gets Malthus going is the disparity of rate of increase of the first necessity, food, as compared with the rate of increase of population, given the second postulate.

“Population, when unchecked,” writes Malthus,

increases in a geometrical ratio. Subsistence increases only in an arithmetical ratio. A slight acquaintance with numbers will shew the immensity of the first power in comparison of the second.

We will return later to Malthus, to see what he does with these postulates in his essay, but let’s use them as an introduction to the more modern vocabulary around this Malthusian viewpoint. The main object to ‘measure’ in terms of how it changes in time, is a *population*, counted as a number of distinct individuals. Malthus’s *geometrical ratio* hypothesis is essentially the assumption is that the *rate of change of the population* at any moment in time, t , is proportional to the size of that population at that moment. This, of course, is but a starting assumption, to be modified with the complexity of the model, which—to be in any serious way realistic— will certainly have to bring in limited food supply as Malthus does, but may bring in other actors (e.g., predators) and other constraints (e.g., disease) as well.

But the naive assumption unadorned with any baroque complications (that the rate of change of the size of the population is just proportional to size of the population,) would predict *exponential growth*—e.g., twice the population twice the rate of growth, etc. Surely an unsustainable state of being, but simple enough.

A mathematical formula that expresses some specific relation between a list of quantities as they vary in time, and their rates of change (the speed in which they increase or decrease in time) is called a *differential equation*. The starting differential equation for Malthus’s *geometrical ratio* hypothesis is easy: Denoting $P(t)$ the size of the population at time t , its ‘rate of change’ (alias: its “derivative with respect to time”) at time t is denoted $\frac{dP}{dt}(t)$. So the constancy of proportionality between population and its rate of change is expressed in the differential equation:

$$(*) \quad \frac{dP}{dt}(t) = \text{some constant} \times P(t).$$

This equation is indeed nothing more than a direct translation of the statement ‘rate of change of population is proportional to size of population.’ Since we haven’t specified what units we use to parametrize time t , and since we haven’t specified the units that describe P as well in cases

where the ‘size of population’ is not given as a discrete number of individuals, but as a somewhat continuous quantity, we haven’t really specified the ‘some constant’ in this equation which will depend on these choices. Moreover, if P is, in fact, a ‘number of individuals’ this equation can be nothing more than shorthand for a Difference Equation approximation to it. That being said, the solution to the differential equation (*) can be expressed as:

$$(**) \quad P(t) = P(0) \cdot 2^T$$

where $P(0)$ is the population size at time 0 (= “NOW”) and T is time proportional to the t of the initial equation (*) and measured in terms of ‘doubling time units’ for the specific population under study (and this is not necessarily well-known units like hours or minutes or seconds). In each unit of doubling-time—in each tick of the T -clock ($T \mapsto T + 1$)—this population will double (according to this model)³.

One way of viewing this Malthusian equation is as not yet a full-fledged *model* but as an ‘opening move’ in what will be a possibly never-finalized attempt to construct a model that reflects all the specific understandings that one achieves as one studies more deeply whatever situation it is that one aims to ‘model.’ An armature, if you wish.

Malthus, himself, takes that approach in his essay. For starters, he emphasizes that limitations of the rate of expansion of food supply goes directly counter to any expectation that in the long term, the ratio dP/dt over P (i.e., rate of change of population compared to size of population) be constant. This gives at least one clear mechanism of *population self-regulation* in the terminology, of some modern writers; see the interesting discussion-article by Peter Turchin, *Does population ecology have general laws?* OIKOS 94 1726. Copenhagen (2001).

Of course, once you decide that the ratio $r := \frac{1}{P}dP/dt$ needn’t be constant, and therefore you’ve opened up the question of its variation (so, $r = r(f_1, f_2, \dots, f_\nu; t)$ can change as a function of time t , and possibly other parameters f_1, f_2, \dots, f_ν) in a sense you’ve kicked the model down the road, because you have no longer made any assertion at all yet about the behavior of population size; all you have done is to have framed an open-ended question: how, and depending on what, does r vary?

Malthus does push this discussion further to describe—among other things—an intrinsic “oscillation,” as he calls it, commenting that it “will not be remarked by superficial observers.” He writes:

We will suppose the means of subsistence in any country just equal to the easy support of its inhabitants. The constant effort towards population, which is found to act even in the most vicious societies, increase the number of people before the means of subsistence are increased. The food therefore which before supported seven millions must now be divided among seven million and a half, or eight million. The poor consequently must live much worse, and many of them be reduced to severe distress. The number of labourers also being above the proportion of the work of the market, the price of labour must tend toward a decrease, while the price of provisions would at the same time tend to rise. The labourer therefore must work harder to earn the same as he did before.

³Malthus estimated that North America—at the time of his writing—was ‘doubling’ its population every 25 years.

During this season of distress, the discouragements to marriage, and the difficulty of rearing a family are so great that population is at a stand. In the mean time the cheapness of labour, the plenty of labourers, and the necessity of an increased industry amongst them, encourage cultivators to employ more labour upon their land, ... till ultimately the means of subsistence become in the same proportion to the population as the period from which we set out. The situation of the labourer being then again tolerably comfortable, the restraints to population are in some degree loosened and the same retrograde and progressive movements with respect to happiness are repeated.

To put Malthus's idea of oscillation into some kind of mathematical vocabulary⁴ let us give ourselves the letters F for food supply, C for cost of provisions, L for number of labourers, W for wages, i.e., price of a labour-hour, and r , as above, for rate of change of population. So, Malthus argues: as P goes up, F goes down essentially linearly per capita, so C goes up causing distress which makes r go down. But even a small exponential is an exponential, so L goes up even though the amount of work necessary doesn't require such a high L causing W to go down, so r goes further down, so P catches up with supplies of provisions forcing a reversal of all the tendencies listed.

If we imagine this made more precise, say with appropriate time-lags and guesses for the general shape of all these dependences, we would be looking at a finite system of linked Difference Equations that would animate this Malthusian oscillation.

Now Malthus's own discussion in the quoted paragraph already provides us with a model featuring a certain type of interlinked dependencies. The Difference Equations I've just alluded to—but haven't, in fact, written down—are not meant to constitute an independent model, but rather to be a faithful translation of Malthus's discussion into mathematical terms, and therefore should be subservient to Malthus's description. If those equations mystify, they will have failed their mission. Moreover, it is legitimate to ask, without prejudice, what—if anything—is the 'value added' in the act of mathematicizing anything? In particular why might we want to provide a mathematical formulation of the paragraph I quoted from Malthus's essay? Here is a tentative list of reasons.

- *Succinctness, possibly*: you would have a linked set of equations encapsulating Malthus's discursive description. A handy mnemonic, even if nothing else.
- *Quantification*: you would be forced to define the variables explicitly, as measurable quantities.
- *The equations can serve as a receptacle*: you might not yet know, or yet want to specify explicitly, the various dependencies listed above, but you might rather wish to allow for some—even if not infinite—flexibility: e.g., r might depend on C and W and even on P , but you might need more data or more experience before you stipulate anything precise about that function $r(C, W, P)$. Even with this type of 'blanks to be filled in later'—e.g., what explicitly is this $r(C, W, P)$?—these equations might well provide a working vocabulary on which to pin whatever you later learn, the dependencies to be specified ever more precisely⁵ as time goes on.

⁴Vaguely analogous to this is the type of oscillation occurring in the solutions to the standard predator-prey (Lotka-Volterra) differential equations.

⁵One danger (or perhaps opportunity) here—once one sets about making guesses regarding the relationship between distinct variables—is the irresistible urge to make over-precise guesses, motivated by convenience, or simplicity rather than experience. *Creative over-precision*, to put a good face on it. This may be very instructive and a good thing to

- *Numerical experimentation*: Once the equations become specific enough you can run computer experiments allowing you to visualize the concrete effect of these interlinked dependencies.
- *Surprise or Confirmation*: When you run these equations numerically, you might be surprised by the outcome, or find that your qualitative expectations are confirmed. But, with any such surprise you certainly can, and possibly should, raise the question: does this surprise point to something legitimate, or is it a warning-signal that my mathematical translation was flawed?
- *The ‘next question’*: You might be led to ask questions on a finer level.

But the main reason to cast it into mathematical terms is to be then able to comprehend it all the better. In any event, obscurity of the mathematical vocabulary is *not* a good reason to believe it more!

That *obscurity* can intensify *emotional impact* is clear in literature and poetry. Nursery rhymes are garlanded with delicious meaninglessnesses and metaphors often overstretch their logic, the very tension of this over-stretching being a source of their power. Consider:

His delights
 were dolphin-like; they showed his back above
 The element he lived in: In his livery
 Walked crowns and crownets; realms and islands
 were
 As plates dropp’d from his pockets.

in Shakespeare’s *Anthony and Cleopatra*.

do, and not at all akin to the mischief of over-precisely calculating quantities to ten decimal places when the margin of error of the calculation would make most of those decimal places meaningless. Here’s an example worth pondering more than is space for in this essay. The totally general—and therefore unusable—predator-prey pair of differential equations has the form

$$\frac{1}{P} \frac{dP}{dt} = f(P, Q); \quad \frac{1}{Q} \frac{dQ}{dt} = g(P, Q)$$

where P and Q are the population sizes of two species that interact with each other, affecting each other’s population growth rate (predator and prey being the eponymous example) and where the functions f and g —as yet unstipulated—describe how these rates are affected. Useless, so far if we are given no hint about what the relations f and g are meant to be. Now the simplest, if not the most realistic guess you might make about these functions is that they be linear functions of P and Q . If you go simpler than that you hardly have posited any interaction at all between your two species. To be sure you would naturally choose appropriate signs for the coefficients of your linear functions in order to model relationships that might be labeled predator/prey. You end up with the classical Lotka-Volterra equations. Are they a realistic model for any actual predator/prey interaction? I wouldn’t know, but I already am happy that—Occam’s razor-style—this system of differential equations stands a chance of being, perhaps, the right start of a discussion; it is an interesting, simple, manageable, toy mathematical machine that might, after being subjected to appropriate modifications dictated by actual experience, approximately model some real-world (predator versus prey) behavior. Here, though, the warning applies: if you are a firm aficionado of these elegant differential equations it is deeply rewarding to study them and their solutions with whatever exactitude you can achieve. Which is a wonderful thing to do in itself. But if you aim to be building some model, unless you also have some control over the stability of the qualitative aspects of your solutions—stability under minor modifications of this idealized model—you run the risk of being a tad too over-precise in your analysis if you dote on the elegant specifics of its solutions. For more on these equations, see: http://www.scholarpedia.org/article/Predator-prey_model.

Incomprehensibility is a central character in the fascinating essay *Über die Unverständlichkeit* (*On the Incomprehensible*) by Friedrich Schlegel published in 1800, and brilliantly discussed in Michel Chaouli's book *The Laboratory of Poetry*⁶. In effect, Schlegel crowns *incomprehensibility* as, hard to believe, the touchstone of inspired meaning: imagine language as a soup, a medium for ideas, and the poet a cook who brings the whole mixture to a boil and who only is certain that the broth is really cooked if the ingredients are so transformed so as to have—in some sense—gotten away from him, and be singed with incomprehensibility.

This almost self-generated change-of-phase of meaning, where the import gets away from the writer in the throes of writing, is perhaps the more noble reasons of incomprehensibility of message. For an example of textual obscurity attributed to quite a different cause, consider this comment, by the later Platonist Ammonius, regarding a certain sentence in the work of Aristotle:

Let us ask why on earth the Philosopher is contented with obscure teaching. We reply that it is just as in the temples, where curtains are used for the purpose of preventing everyone, and especially the impure, from encountering things they are not worthy of meeting. So too Aristotle uses the obscurity of his philosophy as a veil, so that good people may for that reason stretch their minds even more, whereas empty minds that are lost through carelessness will be put to flight by the obscurity when they encounter sentences like these.⁷

That Schlegel's essay comes in tandem with the rise German Romanticism is no wonder. It makes for the beguiling image of the creator bewildered by his own work and necessarily so bewildered! This is but one page from the book of romantic visions of creative lack of control, and loss, as Arthur Sullivan's *The Lost Chord*⁸ is another. But romanticism aside—the Schlegel sentiment resonates with many modern poets. Perfectly typical is the comment of the poet Mark Strand in an interview:

... language takes over, and I follow it. It just sounds right. And I trust the implication of what I'm saying, even though I'm not absolutely sure what it is that I'm saying. I'm just willing to let it be. Because if I were absolutely sure of whatever it was that I said

⁶Reading the essay itself, says Chaouli, is a “confounding experience.”

⁷This is Ammonius (*On Aristotle's Categories* 7.7-14) translated by S. Marc Cohen and Gareth B. Matthews. It occurs in Volume 7, 1991, of the 100-volume series *The Ancient Commentators of Aristotle*, Richard Sorabji (series editor). The quotation above is taken from the review of Volume 99 of that series, *A ton for Aristotle* by David Sedley that appeared in the TLS (June 2013).

⁸I know not what I was playing,
Or what I was dreaming then;
But I struck one chord of music,
Like the sound of a great Amen.
...
I have sought, but I seek it vainly,
That one lost chord divine,
...

in my poems, if I were sure, and could verify it and check it out and feel, yes, I've said what I intended, I don't think the poem would be smarter than I am.

(Mark Strand, *The Art of Poetry No. 77* Interviewed by Wallace Shawn in *The Paris Review* 148 Fall, 1998).

Yes, I agree. But mathematics (alias: that-which-should-be-as-thoroughly-understood-as-humanly-possible⁹) should never gain any further power using *obscurity* as a wedge. Not mathematics, a source of ecstatic clarity of thought!

This is not to say that bafflement isn't the common inner experience, when thinking about math. One must confess: it is. But bafflement, incomprehensibility, can be taken as a signal: there's more work to be done, a deeper layer to uncover.

It is *good* to be confused!

is a marvelous piece of encouragement given by Glenn Stevens to one of his students who came to him, swamped by perplexity when thinking about a certain problem in mathematics.

But I would be glossing over something truly perplexing if I construed Stevens' "It is *good* to be confused!" as merely a statement of encouragement and nothing more, because it is deeper than that. To see this, let's turn this discussion a bit slant and own up to the curious status of mathematics as having, as goal, crystalline transparency; and yet, as often being intrinsically difficult, and perhaps essentially so. Almost ungraspable; obscure, in effect.

It is not that all important mathematics is this way: some of the most crucial insights are immediately graspable, and illuminate far. But our very vocabulary: the *depth of an idea* gives some normative weight to toughness of comprehension. In fact, being led to the brink of incomprehensibility—the limits of knowledge—has held a fascination for mathematicians, and this fascination has led to some of the most important breakthroughs. From the pythagorean concern over the inexpressibility in 'number' of the ratio of the diagonal to the side of a square figure, to more modern issues of *unsolvability* in all its forms, to the *Incompleteness Theorem* of Gödel, to the quest for control of inaccessible cardinals, one feels the attraction of *that which is not known* and possibly will *never be known* at least in terms of the vocabulary of the mathematics of the epoch. Stevens' dictum exhorts one to develop a yen for confronting confusion, as a sign that one may be on an important track: "It is *good* to be confused!"

But even if the best mathematics occurs when confusion is actively sought and wrestled with, we should complain when the opaqueness of a piece of mathematics is shamelessly exploited as a token of rhetorical advantage.

⁹For a description of how one might teach mathematics to children by having genuine conversations with them, see Bob and Ellen Kaplan's *Essence of Math Circle*. See <http://www.themathcircle.org/>