How to Read Mathematics

by

Shai Simonson and Fernando Gouvea

Mathematics is "a language that can neither be read nor understood without initiation." 1

A *reading protocol* is a set of strategies that a reader must use in order to benefit fully from reading the text. Poetry calls for a different set of strategies than fiction, and fiction a different set than non-fiction. It would be ridiculous to *read* fiction and ask oneself what is the author's source for the assertion that the hero is blond and tanned; it would be wrong to *read* non-fiction and not ask such a question. This reading protocol extends to a *viewing* or *listening* protocol in art and music. Indeed, much of the introductory course material in literature, music and art is spent teaching these protocols.

Mathematics has a reading protocol all its own, and just as we learn to read literature, we should learn to read mathematics. Students need to learn how to read mathematics, in the same way they learn how to read a novel or a poem, listen to music, or view a painting. Ed Rothstein's book, *Emblems of Mind*, a fascinating book emphasizing the relationship between mathematics and music, touches implicitly on the reading protocols for mathematics.

When we read a novel we become absorbed in the plot and characters. We try to follow the various plot lines and how each affects the development of the characters. We make sure that the characters become real people to us, both those we admire and those we despise. We do not stop at every word, but imagine the words as brushstrokes in a painting. Even if we are not familiar with a particular word, we can still see the whole picture. We rarely stop to think about individual phrases and sentences. Instead, we let the novel sweep us along with its flow and carry us swiftly to the end. The experience is rewarding, relaxing and thought provoking.

Novelists frequently describe characters by involving them in well-chosen anecdotes, rather than by describing them by well-chosen adjectives. They portray one aspect, then another, then the first again in a new light and so on, as the whole picture grows and comes more and more into focus. This is the way to communicate complex thoughts that defy precise definition.

Mathematical ideas are by nature precise and well defined, so that a precise description is possible in a very short space. Both a mathematics article and a novel are telling a story and developing complex ideas, but a math article does the job with a tiny fraction of the words and symbols of those used in a novel. The beauty in a novel is in the aesthetic way it uses language to evoke emotions and present themes which defy precise definition. The beauty in a mathematics article is in the elegant efficient way it concisely describes precise ideas of great complexity.

What are the common mistakes people make in trying to read mathematics? How can these mistakes be corrected?

Don't Miss the Big Picture

"Reading Mathematics is not at all a linear experience ...Understanding the text requires cross references, scanning, pausing and revisiting" 2

Don't assume that understanding each phrase, will enable you to understand the whole idea. This is like trying to see a portrait painting by staring at each square inch of it from the distance of your nose. You will see the detail, texture and color but miss the portrait completely. A math article tells a story. Try to see what the story is before you delve into the details. You can go in for a closer look once you have built a framework of understanding. Do this just as you might reread a novel.

Don't be a Passive Reader

> *"A three-line proof of a subtle theorem is the distillation of years of activity. Reading mathematics… involves a return to the thinking that went into the writing" 3*

Explore examples for patterns. Try special cases.

A math article usually tells only a small piece of a much larger and longer story. The author usually spends months discovering things, and going down blind alleys. At the end, he organizes it all into a story that covers up all the mistakes (and related motivation), and presents the completed idea in clean neat flow. The way to really understand the idea is to re-create what the author left out. Read between the lines.

Mathematics says a lot with a little. The reader must participate. At every stage, he/she must decide whether or not the idea being presented is clear. Ask yourself these questions:

Why is this idea true?
Do I really believe it?
Could I convince someone else that it is true?
Why didn't the author use a different argument?
Do I have a better argument or method of explaining the idea?
Why didn't the author explain it the way that I understand it?
Is my way wrong?
Do I really get the idea?
Am I missing some subtlety?
Did this author miss a subtlety?
If I can't understand the point, perhaps I can understand a similar but simpler idea?
Which simpler idea?
Is it really necessary to understand this idea?
Can I accept this point without understanding the details of why it is true?
Will my understanding of the whole story suffer from not understanding why the point is true?

Putting too little effort into this participation is like reading a novel without concentrating. After half an hour, you wake up to realize the pages have turned, but you have been daydreaming and don't remember a thing you read.

Don't Read Too Fast

Reading mathematics too quickly results in frustration. A half hour of concentration in a novel might net the average reader 20-60 pages with full comprehension, depending on the novel and the experience of the reader. The same half hour in a math article buys you 0-10 lines depending on the article and how experienced you are at reading mathematics. There is no substitute for work and time. You can speed up your math reading skill by practicing, but be careful. Like any skill, trying too much too fast can set you back and kill your motivation. Imagine trying to do an hour of high-energy aerobics if you have not worked out in two years. You may make it through the first class, but you are not likely to come back. The frustration from seeing the experienced class members effortlessly do twice as much as you, while you moan the whole next day from soreness, is too much to take.

For example, consider the following theorem from Levi Ben Gershon's manuscript Maaseh Hoshev (The Art of Calculation), written in 1321.

"When you add consecutive numbers starting with 1, and the number of numbers you add is odd, the result is equal to the product of the middle number among them times the last number." It is natural for modern day mathematicians to write this as:

$$\sum_{i=1}^{2k+1} i = (k+1)(2k+1)$$

A reader should take as much time to unravel the two-inch version as he would to unravel the two-sentence version. An example of Levi's theorem is that $1 + 2 + 3 + 4 + 5 = 3\times5$.

Make the Idea your Own

The best way to understand what you are reading is to make the idea your own. This means following the idea back to its origin, and rediscovering it for yourself. Mathematicians often say that to understand something you must first *read* it, then write it down in your own words, then teach it to someone else. Everyone has a different set of tools and a different level of "chunking up" complicated ideas. Make the idea fit in with your own perspective and experience.

"When I use a word, it means just what I choose it to mean"

(Humpty Dumpty to Alice in *Through the Looking Glass* by Lewis Carroll)

> *"The meaning is rarely completely transparent, because every symbol or word already represents an extraordinary condensation of concept and reference" 4*

A well-written math text will be careful to use a word in one sense only, making a distinction, say, between *combination* and *permutation* (or *arrangement*). A strict mathematical definition might imply that "yellow rabid dog" and "rabid yellow dog" are different arrangements of words but the same combination of words. Most English speakers would disagree. This extreme precision is utterly foreign to most fiction and poetry writing, where using multiple words, synonyms, and varying descriptions is *de rigueur*.

A reader is expected to know that an *absolute value* is not about some value that happens to be absolute, nor is a *function* about anything functional.

A particular notorious example is the use of "It follows easily that" and equivalent constructs. It means something like this:

One can now check that the next statement is true with a certain amount of essentially mechanical, though perhaps laborious, checking. I, the author, could do it, but it would use up a large amount of space and perhaps not accomplish much, since it'd be best for you to go ahead and do the computation to clarify for yourself what's going on here. I promise that no new ideas are involved, though of course you might need to think a little in order to find just the right combination of good ideas to apply.

In other words, the construct, when used correctly, is a signal to the reader that what's involved here is perhaps tedious and even difficult, but involves no deep insights. The reader is then free to decide whether the level of understanding he/she desires requires going through the details or warrants saying "Okay, I'll accept your word for it."

Now, regardless of your opinion about whether that construct should be used in a particular situation, or whether authors always use it correctly, you should understand what it is supposed to mean. "It follows easily that" does not mean

if you can't see this at once, you're a dope, neither does it mean

this shouldn't take more than two minutes,

but a person who doesn't know the lingo might interpret the phrase in the wrong way, and feel frustrated. This is apart from the issue that one person's tedious task is another person's challenge, so the author must correctly judge the audience.

Know Thyself

Texts are written with a specific audience in mind. Make sure that you are the intended audience, or be willing to do what it takes to become the intended audience.

T.S.Eliot's

A Song for Simeon:

> *Lord, the Roman hyacinths are blooming in bowls and*
> *The winter sun creeps by the snow hills;*
> *The stubborn season has made stand.*
> *My life is light, waiting for the death wind,*
> *Like a feather on the back of my hand.*
> *Dust in sunlight and memory in corners*
> *Wait for the wind that chills towards the dead land.*

For example, Eliot's poem pretty much assumes that its readers are going to either know who Simeon was or be willing to find out. It also assumes that its reader will be somewhat experienced in reading poetry and/or is willing to work to gain such experience. He assumes that they will either know or investigate the allusions here. This goes beyond knowledge of things like who Simeon was. For example, why are the hyacinths "Roman?" Why is that important?

Elliot assumes that the reader will read slowly and pay attention to the images: he juxtaposes dust and memory, relates old age to winter, compares waiting for death with a feather on the back of the hand, etc. He assumes that the reader will recognize this as poetry; in a way, he's assuming that the reader is familiar with a whole poetic tradition. The reader is supposed to notice that alternate lines rhyme, but that the others do not, and so on.

Most of all, he assumes that the reader will read not only with the mind, but also with his/her emotions and imagination, allowing the images to summon up this old man, tired of life but hanging on, waiting expectantly for some crucial event, for something to happen.

Most math books are written with assumptions about the audience: that they know certain things, that they have a certain level of "mathematical maturity," etc. Before you start to read, make sure you know what the author expects you to know.

An Example of Mathematical Writing

To allow an opportunity to experiment with the guidelines presented here, I am including a small piece of mathematics often called the birthday paradox. The first part is a concise mathematical article explaining the problem and solving it. The second is an imaginary Reader's attempt to understand the article by using the appropriate reading protocol. This article's topic is probability and is accessible to a bright and motivated reader with no background at all.

The Birthday Paradox

A professor in a class of 30 random students offers to bet that there are at least two people in the class with the same birthday (month and day, but not necessarily year). Do you accept the bet? What if there were fewer people in the class? Would you bet then?

Assume that the birthdays of n people are uniformly distributed among 365 days of the year (assume no leap years for simplicity). We prove that, the probability that at least two of them have the same birthday (month and day) is equal to:

$$1-\left(\frac{365\times 364\times 363\times\ldots\times(365-n+1)}{365^n}\right)$$

What is the chance that among 30 random people in a room, there are at least two or more with the same birthday? For $n = 30$, the probability of at least one matching birthday is about 71%. This means that with 30 people in your class, the professor should win the bet 71 times out of 100 in the long run. It turns out that with 23 people, she should win about 50% of the time.

Here is the proof: Let P(n) be the probability in question. Let Q(n) = 1 – P(n) be the probability that no two people have a common birthday. Now calculate Q(n) by calculating the number of n birthdays without any duplicates and divide by the total number of n possible birthdays. Then solve for P(n).

The total number of n birthdays without duplicates is:

$$365 \times 364 \times 363 \times \ldots \times (365 - n + 1).$$

This is because there are 365 choices for the first birthday, 364 for the next and so on for n birthdays. The total number of n birthdays without any restriction is just 365^n because there are 365 choices for each of n birthdays. Therefore, Q(n) equals

$$\frac{365\times 364\times 363\times\ldots\times(365-n+1)}{365^n}$$

Solving for P(n) gives P(n) = 1 – Q(n) and hence our result.

Our Reader Attempts to Understand the Birthday Paradox

In this section, a naive Reader tries to make sense out of the last few paragraphs. The Reader's part is a metaphor for the Reader thinking out loud, and the Professional's comments represent research on the Reader's part. The appropriate protocols are centered and bold at various points in the narrative.

My Reader may seem to catch on to things relatively quickly. However, be assured that in reality a great deal of time passes between each of my Reader's comments, and that I have left out many of the Reader's remarks that explore dead-end ideas. To experience what the Reader experiences requires much more than just reading through his/her lines. Think of his/her part as an outline for your own efforts.

Know Thyself

Reader (R): I don't know anything about probability, can I still make it through?

Professional (P): Let's give it a try. We may have to backtrack a lot at each step.

R: What does the phrase "30 random students" mean?

"When I use a word, it means just what I choose it to mean"

P: Good question. It doesn't mean that we have 30 spacy or scatter-brained people. It means we should assume that the birthdays of these 30 people are independent of one another and that every birthday is equally likely for each person. The author writes this more technically a little further on: "Assume that the birthdays of *n* people are uniformly distributed among 365 days of the year."

R: Isn't that obvious? Why bother saying that?

P: Yes the assumption is kind of obvious. The author is just setting the groundwork. The sentence guarantees that everything is normal and the solution does not involve some imaginitive fanciful science-fiction.

R: What do you mean?

P: For example, the author is *not* looking for a solution like this: everyone lives in Independence Land and is born on the 4^{th} of July, so the chance of two or more people with the same birthday is 100%. That is not the kind of solution mathematicians enjoy. Incidentally, the assumption also implies that we do not count leap years. In particular, *nobody* in this problem is born on February 29. Continue reading.

R: I don't understand that long formula, what's *n*?

P: The author is solving the problem for any number of people, not just for 30. The author, from now on, is going to call the number of people *n*.

R: I still don't get it. So what's the answer?

Don't Be a Passive Reader - Try Some Examples

P: Well, if you want the answer for 30, just set $n = 30$.

R: Ok, but that looks complicated to compute. Where's my calculator? Let's see: $365 \times 364 \times 363 \times ... \times 336$. That's tedious, and the final exact value won't even fit on my calculator. It reads:

2.1710301835085570660575334772481e+76

If I can't even calculate the answer once I know the formula, how can I possibly understand where the formula comes from?

P: You are right that this answer is inexact, but if you actually go on and do the division, your answer won't be too far off.

R: The whole thing makes me uncomfortable. I would prefer to be able to calculate it more exactly. Is there another way to do the calculation?

P: How many terms in your product? How many terms in the product on the bottom?

R: You mean 365 is the first term and 364 is the second? Then there are 30 terms. There are also 30 terms on the bottom, (30 copies of 365).

P: Can you calculate the answer now?

R: Oh, I see. I can pair up each top term with each bottom term, and do 365/365 as the first term, then multiply by 364/365, and so on for 30 terms. This way the product never gets too big for my calculator. (After a few minutes)... Okay, I got 0.29368, rounded to 5 places.

P: What does this number mean?

Don't Miss the Big Picture

R: I forgot what I was doing. Let's see. I was calculating the answer for $n = 30$. The 0.29368 is everything except for subtracting from 1. If I keep going I get 0.70632. Now what does that mean?

P: Knowing more about probability would help, but this simply means that the chance that two or more out of the 30 people have the same birthday is 70,632 out of 100,000 or about 71%.

R: That's interesting. I wouldn't have guessed that. You mean that in my class with 30 students, there's a pretty good chance that at least two students have the same birthday?

P: Yes that's right. You might want to take bets before you ask everyone their birthday. Many people don't think that a duplicate will occur. That's why some authors call this the birthday *paradox.*

R: So that's why I should read mathematics, to make a few extra bucks?

P: I see how that might give you some incentive, but I hope the mathematics also inspires you without the monetary prospects.

R: I wonder what the answer is for other values of n. I will try some more calculations.

P: That's a good idea. We can even make a picture out of all your calculations. We could plot a graph of the number of people versus the chance that a duplicate birthday occurs, but maybe this can be left for another time.

R: Oh look, the author did some calculations for me. He says that for $n = 30$ the answer is about 71%; that's what I calculated too. And, for $n = 23$ it's about 50%. Does that make sense? I guess it does. The more people there are, the greater the chance of a common birthday. Hey, I am anticipating the author. Pretty good. Okay, let's go on.

P: Good, now you're telling *me* when to continue.

Don't Read Too Fast

R: It seems that we are up to the proof. This must explain why that formula works. What's this Q(n)? I guess that P stands for probability but what does Q stand for?

P: The author is defining something new. He is using Q just because it's the next letter after P, but Q(n) is also a probability, and closely related to P(n). It's time to take a minute to think. What is Q(n) and why is it equal to 1 – P(n)?

R: Q(n) is the probability that no two people have the same birthday. Why does the author care about that? Don't we want the probability that at least two have the same birthday?

P: Good point. The author doesn't tell you this explicitly, but between the lines, you can infer that he has no clue how to calculate P(n) directly. Instead, he introduces Q(n) which supposedly equals 1 – P(n). Presumably, the author will proceed next to tell us how to compute Q(n). By the way, when you finish this article, you may want to deal with the problem of calculating P(n) directly. That's a perfect follow up to the ideas presented here.

R: First things first.

P: Ok. So once we know Q(n), then what?

R: Then we can get P(n). Because if Q(n) = 1 – P(n), then P(n) = 1 – Q(n). Fine, but why is Q(n) = 1 – P(n)? Does the author assume this is obvious?

P: Yes, he does, but what's worse, he doesn't even tell us that it is obvious. Here's a rule of thumb: when an author says *clearly this is true* or *this is obvious*, then take 15 minutes to convince yourself it is true. If an author doesn't even bother to say this, but just implies it, take a little longer.

R: How will I know when I should stop and think?

P: Just be honest with yourself. When in doubt, stop and think. When too tired, go watch television.

R: So why is Q(n) = 1 – P(n)?

P: Let's imagine a special case. If the chance of getting two or more of the same birthdays is 1/3, then what's the chance of not getting two or more?

R: It's 2/3, because the chance of something not happening is the opposite of the chance of it happening.

Make the Idea Your Own

P: Well, you should be careful when you say things like *opposite*, but you are right. In fact, you have discovered one of the first rules taught in a course on probability. Namely, that the probability that something will not occur is 1 minus the probability that it will occur. Now go on to the next paragraph.

R: It seems to be explaining why Q(n) is equal to long complex-looking formula shown. I will never understand this.

P: The formula for Q(n) is tough to understand and the author is counting on your diligence, persistence, and/or background here to get you through.

R: He seems to be counting all possibilities of something and dividing by the total possibilities, whatever that means. I have no idea why.

P: Maybe I can fill you in here on some background before you try to check out any more details. The probability of the occurrence of a particular type of outcome is defined in mathematics to be: the total number of possible ways that type of outcome can occur divided by the total number of possible outcomes. For example, the probability that you throw a four when throwing a die is 1/6. Because there is one possible 4, and there are six possible outcomes. What's the probability you throw a four or a three?

R: Well I guess 2/6 (or 1/3) because the total number of outcomes is still six but I have two possible outcomes that work.

P: Good. Here's a harder example. What about the chance of throwing a sum of four when you roll two dice? There are three ways to get a four (1-3, 2-2, 3-1) while the total number of possible outcomes is 36. That is 3/36 or 1/12. Look at the following 6 by 6 table and convince yourself.

1-1,	1-2,	1-3,	1-4,	1-5,	1-6
2-1,	2-2,	2-3,	2-4,	2-5,	2-6
3-1,	3-2,	3-3,	3-4,	3-5,	3-6
4-1,	4-2,	4-3,	4-4,	4-5,	4-6
5-1,	5-2,	5-3,	5-4,	5-5,	5-6
6-1,	6-2,	6-3,	6-4,	6-5,	6-6

What about the probability of throwing a 7?

R: Wait. What does 1-1 mean? Doesn't that equal 0?

P: Sorry, my bad. I was using the minus sign as a dash, just to mean a pair of numbers, so 1-1 means a roll of one on each die - snake eyes.

R: Couldn't you have come up with a better notation?

P: Well maybe I could/should have, but commas would look worse, a slash would look like division, and anything else might be just as confusing. We aren't going to publish this transcript anyway.

R: That's a relief. Well, I know what you mean now. To answer your question, I can get a seven in six ways via 1-6, 2-5, 3-4, 4-3, 5-2, or 6-1. The total number of outcomes is still 36, so I get 6/36 or 1/6. That's weird, why isn't the chance of rolling a 4 the same as for rolling a 7?

P: Because not every sum is equally likely. The situation would be very different if we were simply spinning a wheel with the sums 2 through 12 listed in equally spaced intervals. In that case, each one of the 11 sums would have probability 1/11.

R: Okay, now I am an expert. Is probability just about counting?

P: Sometimes, yes. But counting things is not always so easy.

R: I see, let's go on. By the way, did the author really expect me to know all this? My friend took Probability and Statistics and I am not sure he knows all this stuff.

P: There's a lot of information implied in a small bit of mathematics. Yes, the author expected you to know all this, or to discover it yourself just as we have done. If I hadn't been here, you would have had to ask yourself these questions and answer them by thinking, looking in a reference book, or consulting a friend.

R: So the chance that there are no two people with the same birthday is the number of possible sets of n birthdays without a duplicate divided by the total number of possible sets of n birthdays.

P: Excellent summary.

R: I don't like using n, so let me use 30. Perhaps the generalization to n will be easy to see.

P: Great idea. It is often helpful to look at a special case before understanding the general case.

R: So how many sets of 30 birthdays are there total? I can't do it. I guess I need to restrict my view even more. Let's pretend there are only two people.

P: Fine. Now you're thinking like a mathematician. Let's try $n = 2$. How many sets of two birthdays are there total?

R: I number the birthdays from 1 to 365 and forget about leap years. Then these are the all the possibilities:

1-1, 1-2, 1-3, ... , 1-365,

2-1, 2-2, 2-3, ... , 2-365,

...

365-1, 365-2, 365-3, ... , 365-365.

P: When you write 1-1, do you mean 1-1 = 0, as in subtraction?

R: Stop teasing me. You know exactly what I mean.

P: Yes I do, and nice choice of notation I might add. Now how many pairs of birthdays are there?

R: There are 365×365 total possibilities for two people.

P: And how many are there when there are no duplicate birthdays?

R: I can't use 1-1, or 2-2, or 3-3 or ... 365-365, so I get

1-2, 1-3, ... , 1-365,
2-1, 2-3, ... , 2-365,
...
365-1, 365-2, ... , 365-364

The total number here is 365 × 364 since each row now has 364 pairs instead of 365.

P: Good. You are going a little quickly here, but you're 100% right. Can you generalize now to 30? What is the total number of possible sets of 30 birthdays? Take a guess. You're getting good at this.

R: Well if I had to guess, (it's not really a guess, after all, I already know the formula), I would say that for 30 people you get 365 × 365 ×... × 365, 30 times, for the total number of possible sets of birthdays.

P: Exactly. Mathematicians write 365^{30}. And what is the number of possible sets of 30 birthdays without any duplicates?

R: I know the answer *should* be 365 × 364 × 363 × 362 × ... × 336, (that is, start at 365 and multiply by one less for 30 times), but I am not sure I really see why this is true. Perhaps I should do the case with three people first, and work my way up to 30?

P: Splendid idea. Let's quit for today. The whole picture is there for you. When you are rested and you have more time, you can come back and fill in that last bit of understanding.

R: Thanks a lot; it's been an experience. Later.

---

1. *Emblems of Mind*, Edward Rothstein, Avon Books, page 15.
2. ibid, page 16.
3. ibid, page 38
4. ibid, page 16.