

How Craig Barton Wishes he'd Taught Maths

**SIR WILLIAM
TIMOTHY GOWERS**

A couple of months ago, I can't remember precisely how, I became aware of a book called *How I Wish I'd Taught Maths*, by Craig Barton, that seemed to be highly thought of. The basic idea was that Craig Barton is an experienced, and by the sound of things very good, maths teacher who used to take a number of aspects of teaching for granted, until he looked into the mathematics-education literature and came to realize that many of his cherished beliefs were completely wrong. Since I've always been interested in the question of how best to teach mathematics, both because of my own university teaching and because from time to time I like to pontificate about school-level teaching, I decided to order the book. More surprisingly, given my past history of buying books that I felt I ought to read, I read it from cover to cover, all 450 pages of it.

As it happens, the book is ideally designed for people who *don't* necessarily want to read it from cover to cover, because it is arranged as follows. At the top level it is divided into chapters. Each chapter starts with a small introduction and thereafter is divided into sections. And each section has precisely the same organization: it is divided into subsections entitled, "What I used to believe", "Sources of inspiration", "My takeaway", and "What I do now". These are reasonably self-explanatory, but just to spell it out, the first subsection sets out a plausible belief that Craig Barton used to have about good teaching practice, often ending with a rhetorical question such as "What could possibly be wrong with that?", the second is a list of references (none of which I have yet followed up, but some of them look very interesting), the third is a discussion of what he learned from the references, and the last one is about how he put that into practice. Also, each chapter

Keywords: Mathematics, pedagogy, beliefs, unlearning, learning

ends with a short subsection entitled “If I only remember three things ...”, where he gives three sentences that sum up what he thinks is most important in the chapter.

One question I had in the back of my mind when reading the book was whether any of it applied to teaching at university level. I’m still not sure what I think about that. There is a reason to think not, because the focus of the book is very much on school-level teaching, and many of the challenges that arise do not have obvious analogues at university level. For example, he mentioned (on page 235) the following fascinating experiment, where people were asked to do the following multiple-choice question and then justify their answers.

Which of these values could not represent a probability?

- A. $2/3$
- B. 0.72315
- C. 1.46
- D. 0.002

Let me quote the book itself for a discussion of this question.

*Surely the rule *probabilities must be less than or equal to 1* is about as straightforward as it gets in maths? But why, then, did 47% of the 5000+ students who answered this question get it wrong?*

A few students’ explanations reveal all:

I think B because it’s just a massive decimal and the rest look pretty legit. I also don’t see how a number that big could be correct.

I think B because you wouldn’t see this in probability questions.

I think D because you can’t have 0.002 as an answer because it is too low.

If students are only used to meeting ‘nice-looking’ probabilities during examples and practice questions, then it is little surprise they come a cropper when they encounter strange-looking answers.

Could one devise a university-level question that would catch a significant proportion of people out in a similar way? I’m not sure, but here’s an attempt.

Which of the following is not a vector space with the obvious notions of addition and scalar multiplication?

- A. The set of all complex numbers.
- B. The set of all functions from $(0, 1)$ to \mathbb{R} that are twice differentiable.
- C. The set of all polynomials in x with real coefficients that have $x^2 + x + 1$ as a factor.
- D. The set of all triples (a, b, c) of integers.
- E. The set of all sequences $(x_1, \dots, x_n) \in \mathbb{R}^n$ such that $x_1 + \dots + x_n = 0$ and $x_1 + 2x_2 + \dots + nx_n = 0$.

I think at Cambridge almost everyone would get this question right (though I’d love to do the experiment). But Cambridge mathematics undergraduates have been selected specifically to study mathematics. Perhaps at a US university, before people have chosen their majors, people might be tempted to choose another option (such as B, because vector spaces are to do with algebra and not calculus), while not noting that the obvious scalars in D do not form a field. Or perhaps they wouldn’t like A because the scalar field is the same as the set of vectors (unless, that is, they thought that the obvious scalars were the real numbers).

More generally, I feel that there are certain kinds of mistakes that are commonly made at school level that are much less common at university level simply because those who survive long enough to reach that stage have been trained not to make them. For example, at university level we become used to formal definitions. Once one is in the habit of using these, deciding whether a structure is a vector space is simply a question of seeing whether the definition of a vector space applies, rather than thinking “Hmm, does that look like the vector spaces I’ve met up to now?” We also become part of a culture where it is common to look at pathological, or at least slightly surprising, examples of concepts, and so on.

Another reason I decided to read the book was that I have certain prejudices about the teaching of mathematics at school level and I was interested to know whether they would be reinforced by the book or challenged by it. This was a win-win situation, since it is always nice to have one's prejudices confirmed, but also rather exhilarating to find out that something that seems obviously correct is in fact wrong.

A prejudice that was strongly confirmed was the value of mathematical fluency. Barton says, and I agree with him (and suggested something like it in my book *Mathematics, A Very Short Introduction*) that it is often a good idea to teach fluency first and understanding later. More precisely, in order to decide whether it is a good idea, one should assess (i) how difficult it is to give an explanation of why some procedure works and (ii) how difficult it is to learn how to apply the procedure without understanding why it works.

For instance, suppose you want to teach multiplication of negative numbers. The rule "If they have the same sign then the answer is positive, and if they have different signs then the answer is negative" is a short and straightforward rule, but explaining why -2 times -3 should equal 6 is not very straightforward. So if one begins with the explanation, there is a big risk of conveying the idea that multiplication of negative numbers is a difficult, complicated topic, whereas if one gives plenty of practice in applying the simple rules, then one gives one's students fluency in an operation that comes up in many other contexts (such as, for instance, multiplying $(x - 2)$ by $(x - 3)$), and one can try to justify the rule later, when they are comfortable with the rule itself. I remember enjoying the challenge of thinking about why the rule for dividing one fraction by another was correct, but that was long after I was happy with using the rule itself. I don't remember being bothered by the lack of justification up to that point.

As an example in the other direction, Barton gives that of solving linear equations. The danger here is that one can learn a procedure for solving equations such as $2x + 3 = 17$, get good at it, and

then be completely stuck when faced with an equation such as $4 - 2x = 3x - 11$. Here a bit of understanding can greatly help. Barton advocates something called the balance method, where one imagines both sides of the equation on a balance, and one is required to make sure that balance is maintained the whole time. I think (but without too much confidence after reading this book) that I would go for something roughly equivalent, but not quite the same, which is to stress the rule *you can do the same thing to both sides of an equation* (worrying about things like squaring both sides or multiplying by zero later). Then the problem of solving linear equations would be reduced to a kind of puzzle: what can we do to both sides of this equation to make the whole thing look simpler?

That last question is related to another fascinating nugget that is mentioned in the book. Barton gives an example of a question concerning a parallelogram ABCD, where the angle at A is 105 degrees. The line BC is extended to a point E, which is then joined by an additional line segment to D, and the angle CED is 30 degrees. The question is to prove that the triangle CED is isosceles.

Apparently, this question is found hard, because one cannot achieve the goal in one step. Instead, one must observe that the angle of the parallelogram at C is also 105 degrees, from which it follows that the angle ECD is 75 degrees. And from that it follows that the angle EDC is 75 degrees as well, and the problem is solved.

But the interesting thing is that if you change the question to the more open-ended question, "Fill in as many angles in this diagram as you can," then many people who found the goal-oriented version too hard have no difficulty in filling out all the angles in the diagram and therefore noticing that the triangle CED is isosceles.

The lesson I would draw from this with the equations question is that instead of asking for a solution to the equation $4 - 2x = 3x - 11$, it might be better to ask "See whether you can make the equation look simpler by doing

something to both sides. If you manage, see if you can then make it even simpler. Keep going until you have made it as simple as you can.” This would of course come after they had already seen several examples of the kind of thing one can do to both sides of an equation.

Barton isn't content with just telling the reader that certain methods of teaching are better than others: he also tells us the theory behind them. Of particular importance, he claims, is the fact that we cannot hold very much in our short-term memory. This was music to my ears, as it has long been a belief of mine that the limited capacity of our short-term memory is a hugely important part of the answer to the question of why mathematics looks as it does, by which I mean why, out of all the well-formed mathematical statements one could produce, the ones we find interesting are those particular ones. I have even written about this (in an article entitled Mathematics, Memory and Mental Arithmetic, which unfortunately appeared in a book and is not available online, but I might try to do something about that at some point).

This basic point informs a lot of the discussion in the book. Consider, for example, a question that asked you to find the perimeter of a rectangle that had side lengths $\frac{2}{3}$ and $\frac{3}{5}$. This could be a great question, but it is very important to ask it at the right point in the students' development. If you ask it before they are fluent at adding fractions and at working out perimeters of rectangles, then the amount they have to hold in their heads may well exceed their cognitive capacity: they need to store the fact that you have to add the two lengths, and multiply by 2, and put both fractions over a common denominator. It is to avoid this kind of strain that attaining fluency is so important: it literally makes it easier to think, and in particular to solve the kind of interesting problems we would all like them to be able to solve. Barton absolutely doesn't dispute the value of interesting problems that mix different parts of mathematics — he just argues, very convincingly, that one has to be careful when to introduce them.

An idea he discusses a lot, and that I think might perhaps have a role to play in university-level teaching, is what he calls diagnostic questions, and in particular low-stakes diagnostic tests. These typically take the form of a short multiple-choice quiz, and he tries very hard to create a classroom culture where people understand that the purpose of the quiz is not assessment — the quizzes do not “count” for anything — but a tool to help learning, and in particular to help diagnose problems with understanding.

What makes these questions “diagnostic” is that they are carefully designed in such a way that if you have a certain misconception, then you will be drawn towards a certain wrong answer. That is, the wrong answers people give are informative for the teacher, rather than merely wrong. Here, for example, is a question that fails to be diagnostic followed by a modified version that succeeds.

A triangle has one side of length 6 and two sides of length 5. What is its area?

- A. 8
- B. 11
- C. 12
- D. 15
- E. 20

- A. 6
- B. 12
- C. 15
- D. 16
- E. 24
- F. 30

With the second set of choices, each answer has a potential route that one can imagine somebody taking. To obtain the answer 6, one chops the triangle into two right-angled triangles, each of height 4 and base 3, calculates the area of one of them, and forgets to double it. The correct answer is 12. To obtain 15, one takes the formula “half the height times the base” but substitutes in 5 for the height. To obtain 16 one calculates the perimeter. To obtain the answer 24 one takes the height times the base.

And to obtain the answer 30 one multiplies the two numbers 6 and 5 together (on the grounds that “to calculate the area you multiply the two numbers together”). Thus, wrong answers yield useful information. With the first set of answers, that just isn’t the case — they are much more likely to be the result of pure guesswork.

It’s worth mentioning that Terence Tao has created a number of multiple-choice quizzes (<http://scherk.pbworks.com/w/page/14864181/FrontPage>) on university-level topics. He has also blogged about it here (<https://terrytao.wordpress.com/2008/12/14/on-multiple-choice-questions-in-mathematics/>). They are not exactly diagnostic in the sense Barton is talking about, but one could imagine trying to make them so.

Barton uses these diagnostic tests to get a much clearer picture of what his class already understands, before he launches into the discussion of some new topic, than he would by simply asking questions to the class and getting answers from a few keen students. If he diagnoses a fairly serious collective misunderstanding, then he will spend time dealing with that, rather than pointlessly trying to build on shaky foundations.

I’m jumping around a bit here, but a semi-counterintuitive idea that he advocates, which is apparently backed up by serious research, is what he calls pretesting. This means testing people on material that they have not yet been taught. As long as this is done carefully, so that it doesn’t put students off completely, this turns out to be very valuable, because it prepares the brain to be receptive to the idea that will help to solve that pesky problem. And indeed, after a moment of getting used to the idea, I found it not counterintuitive at all. In fact, it resonates very strongly with my experience as a research mathematician: I find reading other people’s papers very difficult as a rule, but if they can help me solve a problem I’m working on, a lot of that difficulty seems to melt away, because I know exactly what I want, and am looking out for the key idea that will give it to me.

There’s a great section on the use of artificial “real-world” problems. I think he would agree with me about Use of Maths A-level (<https://gowers.wordpress.com/2009/07/11/help-im-stuck-in-my-ivorytower/>). As someone he quotes says, “Students are constantly on their guard against being conned into being interested.” An example he discusses is

Alan drinks $5/8$ of a pint of beer. What fraction of his drink is left?

If the entire point of the exercise is to gain fluency with subtracting fractions, then he advocates just cutting the crap and asking them to calculate $1-5/8$, which I agree with 100%.

If, on the other hand, it is intended as an exercise in stripping away the unnecessary real-world stuff and getting at the underlying mathematics, then he has interesting things to say (later in the book than this section) about the relationship between what he calls the surface structure and the deep structure. The former is to do with the elements of the question that present themselves directly to the student — in this case Alan and the beer — while the deep structure is more like the underlying mathematical question. To train people to uncover the deep structure, it is very important to give them pairs of questions with the same surface structure and different deep structures, and vice versa. Otherwise, they may learn a procedure that works for lots of similar examples and lets them down as soon as a new example comes along with a different deep structure.

There is lots more in the book — obviously, given its length — but I hope this conveys some of its flavour. The only negative thing I can think of to say is that the word “flipping” is overused — the sentence “Teaching is flipping hard” occurs several times, when once would be enough for one book. But if you’re ready for a bit of jocular provocation of that kind, then I recommend it, as I found it highly thought provoking. I don’t yet know what the result of that provocation will be, but I’m pretty sure there will be one.

This review has been reprinted with kind permission from Prof Timothy Gowers. The original review was published in Prof Gowers's Weblog, <https://gowers.wordpress.com/2018/12/22/how-craig-barton-wishes-hed-taught-maths/>.



SIR WILLIAM TIMOTHY GOWERS, FRS is a Royal Society Research Professor at the Department of Pure Mathematics and Mathematical Statistics at the University of Cambridge, where he also holds the Rouse Ball chair, and is a Fellow of Trinity College, Cambridge. In 1998, he received the Fields Medal for research connecting the fields of functional analysis and combinatorics. Prof Gowers is deeply interested in mathematics education and his thoughts and insights might be found at his weblog (<https://gowers.wordpress.com/>).