Questions Pupils Ask!

Why isn't 1 a prime number?

by **Colin Foster**

"My old teacher said that 1 was a prime number."

I have often been told this, and I have little doubt that most of these 'old teachers' never said any such thing! (When my pupils move on from me, I dread to think what they tell their new teachers that they think I said!) Nevertheless, wherever it came from, the belief that 1 is a prime number comes up quite a lot in mathematics classrooms. In a way, it is quite a reasonable thing to say – and may indeed follow logically from pupils' definitions of prime. Definitions vary, but a common one is: 'A prime number is a number whose only factors are 1 and itself.' The phrase '1 and itself' seems to be the memorable part, but it is problematic, because what if 'itself' is 1? Do the '1' and 'itself' have to be distinct or not? These definitions seem to be ambiguous regarding whether 1 should be called a prime number or not, which is probably one reason why there is a bit of confusion about it.

There are various possible explanations given in mathematics classrooms for why 1 isn't prime. One approach is to regard the number 1 as an exception – mathematical statements have exceptions too, and exceptions at the beginning of a sequence are somehow a bit more respectable than exceptions that occur some way further down. So we point out that 1 is a funny number because $1 \times 1 = 1$, and we designate it a special case. So we could say: 'A prime number is a number *greater than 1* whose only factors are 1 and itself.' But this feels like a bit of a fudge, and the 'greater than 1' bit is quite likely to be forgotten unless there is some discussion of why we need to include it. And is it really fair to say that 1 is an exception in this context? After all, the number 2 is also a kind of exception, being the only *even* prime, and some theorems (for example, 'All prime numbers are odd') are valid 'for all primes greater than 2'. Indeed, the Sieve of Eratosthenes is based on the idea that every prime p is a unique exception to the statement that 'There are no multiples of p among the primes' (Note 1). (Of course, you could say that the problem here is that we count $1 \times p$ as a multiple of p.) We happen to have the word 'even' for multiples of 2, but if we had a word 'threeven', let's say, for multiples of 3, then we would be able to say 'All prime numbers greater than three are not threeven'. So perhaps 1 isn't any more special than any of the primes.

Another justification that I have heard is that '1 is a square number, and prime numbers are never square, so 1 can't be a prime number.' What do you think of that? It makes me wonder: How do we know that prime numbers are never square? It seems to beg the question whether 1 is prime: certainly prime numbers *greater than 1* are never square! Another common resolution is: 'A prime number is a number with exactly two (distinct) factors. The number 1 has only one factor, so doesn't have enough factors to be a prime number.' This seems quite clear, and an improvement on the '1 and itself' family of definitions, but why should we be so interested in numbers with exactly two factors that we give them a special name? Why are they more important than, say, numbers with exactly three factors? Numbers with exactly three factors must be squares of primes, but as far as I know they don't have a special name (Foster, 2016).

Perhaps excluding 1 from the primes is simply an example of an arbitrary convention? The community of mathematicians collectively decides not to call 1 prime, and in another universe they might have done otherwise, so we just need to accept it and move on. Indeed, historically many mathematicians up to the nineteenth century thought of 1 as prime – Henri Lebesgue (1875–1941) is usually said to have been the last professional mathematician to call 1 prime. (The Greeks didn't regard 1 as prime, but that's because they didn't regard it as a number at all!)

For me, none of these responses quite gets to the heart of why prime numbers are important and why it might be a good idea to exclude 1 from them. The importance of prime numbers arises from the *fundamental theorem of arithmetic*, the fact that every integer can be expressed uniquely (apart from order) as a product of primes. This makes primes the multiplicative building blocks of the natural numbers, in a similar way to how in chemistry the elements are the building blocks of all the molecules, or the 26 letters of our alphabet are the building blocks of all the words (Note 2). So the problem with 1 is that if we were to include it among the primes it would destroy our lovely theorem, because $24 = 1 \times 2 \times 2 \times 2 \times 3 = 1 \times 1 \times 1 \times 2 \times 2 \times 2 \times 3$, etc. Unique factorization is very useful, so we choose to exclude 1 from the primes because we would rather have unique factorization and have to make an exception for 1 than keep 1 in with primes but have to throw out the property of unique factorization. So, looked at this way, the idea that 1 isn't prime is certainly a *choice* – but a very natural and sensible choice. Since it is a choice, we shouldn't imply to pupils that it is obvious, and we shouldn't make them feel foolish if they struggle with this and initially find it counter-intuitive.

Notes

1. For a lesson introducing prime numbers using the Sieve of Eratosthenes, see Foster (2015).
2. There are really nice visualizations of prime numbers at **www.datapointed.net/visualizations/math/factorization/animated-diagrams/** and **www.ptolemy.co.uk/wp-content/uploads/2009/08/primitives.swf**.

References

Foster, C. 2015 'Prime Suspects', *Teach Secondary*, **4**, 2, pp. 41–43.
Foster, C. 2016 'Thirty Factors', *Mathematics in School*, **45**, 2, pp. 25–27.

Keywords: Definitions; Prime numbers; Unique factorization.

Author Colin Foster, School of Education, University of Nottingham, Jubilee Campus, Wollaton Road, Nottingham NG8 1BB.
e-mail: **colin@foster77.co.uk**
website: **www.foster77.co.uk**

• LETTER • LETTER • LETTER • LETTER • LETTER • LETTER •

Reprinted Letter and Apology
The March issue carried two letters on the subject of Cantorian set theory. The letter from Gerry Leversha contained a double-printing of the set {1, 2} (page 34, about 2/3 of the way down column 3). It should have appeared just once. The way in which Richard Bridge's letter was set was fundamentally flawed, with expressions and formulas appearing in the wrong place. The Editors apologise unreservedly to Gerry and to Richard, and reprint Richard's letter below.

Dear Editors

David Womack seems to reject the significance of Cantor's discovery of a hierarchy of infinite cardinal numbers – or perhaps even to regard Cantor's work as incorrect. However, he appears to base his arguments on a distinction between 'complete' and 'ongoing' 1–1 correspondences which I believe is fallacious. The 1–1 correspondence between the counting numbers and the list of rationals in his Figure 1 is 'complete' in the sense that there is a function which describes it. A rational a/b ($a < b$) will be found opposite the natural number

$$a+\frac{(b-1)(b-2)}{2}$$

(and the mapping may be inverted with slightly greater difficulty).

Similarly, Womack's Figure 2 lists the set of terminating decimal fractions in a way which demonstrates their countability. The number $c/10^d$ appears at position

$$\frac{10}{9}\left(10^{d-1}-1\right)-d+1+c.$$

However, this list excludes any rational such as 1/3 whose decimal expansion recurs, let alone any irrationals.

Womack claims, correctly, that either list may be 'supplemented' by a number that is missing. If we think in more detail about how this supplementation is to take place, we uncover the problem with his later argument. Let's call the function for the correspondence in Figure 2

$$n_0(c,d)\equiv\frac{10}{9}\left(10^{d-1}-1\right)-d+1+c.$$

The list (of all terminating decimals) could be 'supplemented' by a missing number (such as 1/3) by adding it at the start. This would require a new 1–1 function $n_1(c, d) = n_0(c, d) + 1$ [together with $n_1(1/3) = 1$]. However, it is not possible to add 1/3 at the **end** of the list of terminating decimals, as this would require $n_1(1/3) = \aleph_0 + 1$. ($\aleph_0$, 'aleph null', is the conventional symbol for the countable infinity.) This is not a valid expression for a normal 1–1 function, whether 'ongoing' or not. A **finite** number m of numbers could be added at the start in the same way, by letting $n_1(c, d) = n_0(c, d) + m$. However, it is not possible to add an **infinite** number of numbers at the start, as this would require

$$n_1(c, d) = n_0(c, d) + \aleph_0.$$

Again, this is not a valid 1–1 function.

Even so, there is a way of adding even a countably infinite number of numbers to an existing countable list. Let's call our 'staircase' function for the rationals

$$m_0(a,b)\equiv a+\frac{(b-1)(b-2)}{2}.$$

The idea is to shift all the terminating decimals up to the even-numbered positions, opening up a (countably) infinite number of gaps at the odd-numbered positions. The rationals go into these. The required function is:

$$n_1(a, b \text{ representing } a/b) = 2m_0(a, b) - 1$$

and

$$n_1(c, d \text{ representing } c/10^d) = 2n_0(c, d).$$

This method is sometimes known as Hilbert's Hotel. It's not possible to add an **uncountably** infinite number of numbers this way, however, as no 1–1 function like m_0 mapping them to the natural numbers can exist (this is what Cantor's famous Diagonal Proof shows).

In summary, Womack's list of terminating decimals is countable and excludes uncountably many real numbers. It is not possible to 'supplement' it with all of these while leaving it countable, as Womack appears to claim. We **do** need Cantor's wonderful gift of multiple infinities to do full justice to the set of real numbers and the concept of the infinite.

Richard Bridges