Confidence and 'negative' marking

Colin Foster describes a practical way of assessing pupils' confidence in their mathematical answers.

'Negative' marking, in which marks are *deducted* for wrong answers, may strike you as a bad idea. Why would you want to inflict something 'negative' on pupils (Foster, 2007)? It is often said that tests should be about being rewarded for what you know and can do, not being punished for your mistakes. So it is hard to see how 'negative marking' could contribute to a positive classroom ethos. Surely this would just be discouraging for pupils?

The problem with positive marking is that it rewards guessing, which can have some undesirable consequences. In high-stakes school assessments pupils may be told to guess if they are not sure. Teachers may say, "There's nothing worse than leaving a blank answer. That will definitely be marked wrong! Always put something, even if it's a complete guess. It might be right!" But if pupils carry this advice into the classroom and react to uncertainty by guessing, this can create problems. I once observed a pair of pupils working through a sheet of 10 questions, thinking hard about them and getting the correct answers to the first few. After they had completed the first four the teacher said to the class: "You've got 10 more seconds to finish off." One of the pupils said, "Quick, just put anything for the last ones," and the other pupil quickly wrote in some numbers. The class then marked their answers and this pair got 4/10, the first four correct and the last six wrong. I wondered how the teacher would interpret this when they collected in the sheets. Would they realise what had happened or would they spend time wondering what made the last six questions harder than the first four? Might they try to work out how the pupils had obtained those wrong answers and look for what misconceptions might lay behind them?

There may be times in a mathematics lesson when you might want pupils to use their intuition and guess or 'guesstimate' an answer. But this would be a starting point rather than an end point. You might ask them to conjecture something and then work on reasoning it out and deciding whether they needed to modify their initial conjecture. But a guess as a final answer is not helpful to anyone, especially if there is no acknowledgement that it is just a guess. In the case described above, it would have been far more useful for the teacher if the students had simply left the last six questions blank. Guessing in formative assessments just adds statistical noise.

Confidence assessment

It may be helpful to the teacher to know how strongly a pupil believes in the answer that they are giving. But confidence at this fine-grained level is difficult to assess reliably, since if pupils are asked to put up their hands or use traffic lights to indicate their confidence they may be inclined to exaggerate. 'Negative' marking discourages errors but can also prevent pupils from having a go. If pupils are too risk-averse they simply do not get a chance to show what they can do. What is needed is a way for pupils to indicate how sure they are of their answers and to be rewarded for accurate assessment of their own performance.

Confidence assessment is a way to do this. Pupils place a mark from 0 to 10 on each answer that they give. If they are correct they obtain that number of marks. If they are incorrect then they *lose* that number of marks from their total. This approach incentivises pupils to give truthful confidence ratings, because on average you will not benefit from systematically underestimating or overestimating your confidence. It seems possible that by using confidence assessment over a period of time pupils could become more self-aware and teachers might gain insight into how sure their pupils are about the answers that they are giving.

I recently carried out a study (Foster, 2016) exploring using confidence assessment to support pupils' learning of directed (positive and negative) numbers. I chose directed numbers as this allowed the teachers to have a discussion with the pupils beforehand about the mathematics of the process of 'negative marking'. The study took place in five secondary schools located in three different cities in the UK. 345 pupils aged 11-14 in 14 classes across the schools were given 10 questions on directed numbers. They were asked to state alongside each of their answers a confidence level from 0 to 10 expressing how certain they were that each of their answers was correct. They were

told that their total mark would be the sum of the confidence ratings for the correct responses *minus* the sum of the confidence ratings for the *incorrect* responses. The teacher led a short discussion before they started to help the pupils understand the process, asking questions such as "What's the highest possible mark you could get on this sheet? What's the lowest?" Pupils generally realised quite quickly that they could get a maximum of 100 by getting every question right and putting a confidence rating of 10 for each answer, and a minimum of −100 by getting every question wrong and putting a 10 for each answer. Pupils marked their own answers and calculated their own scores. They were then asked to write down their views on the idea of assessing, and rewarding, confidence in this way.

Results

The pupils' facility (number of correct answers out of 10) correlated with their mean confidence ratings with a Pearson's correlation coefficient[1] of 0.546, indicating that they were generally 'well calibrated'. This implies that the pupils understood the confidence assessment process and were not just putting down random numbers for the confidence ratings, suggesting that the confidence data is meaningful and potentially useful information for the teacher. I tested to see whether boys and girls displayed different relationships between facility and confidence and found that girls demonstrated lower confidence than their facility would justify whereas boys demonstrated higher confidence than warranted by their facility. The relative overconfidence of boys is consistent with previous findings (see, for example, Frost, Hyde and Fennema, 1994).

In their comments, the vast majority of the pupils were positive about the confidence assessment approach, despite the differences from more usual assessment practices in schools. There was no significant difference in the total score (on the −100 to 100 scale) of those expressing positive and negative views, indicating that the pupils were not just stating a self-interested preference based on their success. Some pupils volunteered that they would like to use the approach again, saying, for example, "I really like the how sure ... [a]re you process. I would like to use this more often in class". Sometimes this was because of, rather than despite, its difficulty; for example, "I found this challenging and I think you should do this every lesson". The vast majority expressed no concerns regarding the fairness of the approach.

Many pupils said that they were surprised by their score, and some felt that it had made them think more deeply, saying, for example, "I think the 'How Sure' column makes you think more". Some felt that it could increase their confidence. For example, "I think it's a Good [sic] idea because it incourges [sic] people to believe in them selfs [sic]". One pupil commented, "I think that this was a good idea because most of the time kids lie and just guess so this is a good process". Another said, "I think it's good to find out how confident people are with their answer because you might guess (and not feel confident) and get it right. This tells the teacher that you're comfortable, when you're not".

Conclusion

Part of becoming an expert in any area is gaining a better awareness of what you know and also what you *do not* yet know in that area. Knowing your own capabilities is extremely valuable. Felix Okoye is often quoted as saying: "It would be better not to know so many things than to know so many things that are not so." If you know what you do not know, you can focus more on it, seek help, look things up or use a computer or calculator. Even pupils who are generally successful in mathematics are often not sure what they know and what they do not know. I have seen children receiving back a test in which they have scored 100% and saying: "No way! I can't believe I got them all right!" Maybe they are just being modest, but if not then it suggests that although they are competent they do not have too accurate a sense of their capabilities.

In life we can see the damage caused by over-confidence (the banking crisis, for example). Charles Darwin (1871) went as far as to say, "Ignorance more frequently begets confidence than does knowledge" (p. 3). I would not want my doctor to guess when they are unsure what medication to give me. I would rather they look it up and get it right. This is why medicine is one of the areas where students do frequently take tests which are negatively marked. We do not want to simply ignore it when they get questions wrong. But perhaps the same is true for anything that really matters? In an information-rich world, looking things up is often easy; knowing *when* you need to look things up and when you do not is the hard bit.

It might be feared that this kind of confidence assessment could weaken pupils' self-confidence. I would argue that this is *desirable* when that confidence is misplaced. Confidence assessment attempts to increase *appropriate* levels of self-confidence and support *realistic* self-awareness in

order to aid future growth. Using this approach over time, pupils' confidence should self-correct, because they cannot achieve consistently well by guessing or by repeatedly over- or under-reporting their confidence levels. Indeed, studies with university students have found that asking them how sure they are about an answer prompts them to question why they believe it to be correct, leading to self-checking, self-explanation and higher-level reasoning.

Gattegno (1987) drew attention to the difference between a pupil answering the question "2 + 3" with a querying intonation "Five?" as opposed to a more declamatory "Five!", interpreting these as indicators of differing levels of confidence that could benefit from quite different teacher responses. Holt (1990) vividly described 'guess-and-hope' strategies that pupils resorted to whenever they were rewarded for it. Confidence assessment, incorporating 'negative' marking, seems to provide a way of rewarding and encouraging realistic self-confidence in mathematics.

Colin Foster is an Assistant Professor in the School of Education at the University of Nottingham

Note
[1]Pearson's product-moment correlation coefficient is a measure of the linear correlation between two variables. It takes values between +1 (perfect positive correlation) and −1 (perfect negative correlation), where zero indicates no linear correlation at all.

References
Darwin, C. (1871). *The Descent of Man and Selection In Relation To Sex*. Vol. I. London: John Murray.

Foster, C. (2007, May 11). There's a positive side to negative marking. *Times Educational Supplement*, p. 27.

Foster, C. (2016). Confidence and competence with mathematical procedures. *Educational Studies in Mathematics*, *91*(2), 271–288. Freely available with open access at: http://link.springer.com/article/10.1007/s10649-015-9660-9.

Frost, L.A., Hyde, J.S., & Fennema, E. (1994). Gender, mathematics performance, and mathematics-related attitudes and affect: a meta-analytic synthesis. *International Journal of Educational Research*, *21*(4), 373–385.

Gattegno, C. (1987). *Science of Education.* New York: Educational Solutions.

Holt, J. (1990). *How Children Fail*. London: Penguin.

Directed Numbers

For each of the 10 questions below, write the answer in the "Answer" box.

Each time, also write in the "How Sure?" box a whole number from 0 to 10 to indicate how sure you are that your answer is correct. On this scale 0 is "completely unsure" and 10 is "completely certain".

	Question	Answer	Leave Blank	How Sure?
1	5 − 8			
2	− 2 + 7			
3	− 3 − 5			
4	6 − (− 1)			
5	4 + (−9)			
6	(−3) × 4			
7	5 × (−2)			
8	(−2) × (−3)			
9	$\frac{15}{-3}$			
10	$\frac{-10}{-2}$			
Comments				

Figure 1. Pupil task sheet. When marking, pupils enter + or – in the "Leave Blank" column, according to whether the answer is correct or incorrect, and then sum the signed "How Sure?" numbers to give the total mark.