NEW ZEALAND COUNCIL FOR EDUCATIONAL RESEARCH
TE RŪNANGA O AOTEAROA MŌ TE RANGAHAU I TE MĀTAURANGA

National standards for mathematics: where to set the bar?

Some months on from the introduction of National Standards, we now have data about how student performance in a common assessment tool used by schools, NZCER's Progressive Achievement Test (PAT) in mathematics, equates with judgments made using the new National Standards for Mathematics. The information comes from a script scrutiny exercise carried out by the Ministry of Education and it contains a challenge.

The challenge is this: many students whom the PAT would show as making "normal" progress for their year level—not brilliant but not behind either—would be judged to be "failing" when assessed against their Year level National Standard for Mathematics. For example, at Year 8 results from the script scrutiny exercise would indicate that only about 35 percent of students perform at levels at or above those required by the National Standards. What should we take from this seeming mismatch? There are several possible explanations, each of which we now discuss. This exploration then leads us to a complex question that we believe needs a much wider public discussion.

1. Could the PAT norms be wrong?

PATs are carefully developed to identify "normal" progress for our New Zealand children. Any child can be compared with their age cohort to determine their progress relative to others. The norms were determined by a large-scale trial with nationally representative groups of students and statistical analysis of the results. We are confident that they are a robust reflection of progress in the areas of mathematics deemed by the mathematics education community to be important for present and future learning in the subject.

We could argue that the PAT norms might become out of date if our students change the rate at which they make progress. Perhaps they have improved rapidly? However NZCER runs a marking service for PATs which generates large volumes of new data every year and these are used to check that the norms are still correct and current. So we can rule out this possibility.

If we accept that the PAT norms are accurate, we could argue that normal progress is not good enough. In this case, the gap between the norms and the National Standard at the same level would exist because there is a national problem with our children's progress in mathematics. This would certainly be a serious challenge for our education system and we will come back to this question in the last section of the paper.

2. Were the script scrutiny assessors too "tough" when judging the PAT scripts against the standard?

The PAT, like asTTle, another widely used assessment tool, makes a quantitative judgment. Students get individual items right or wrong. The number of the items they get correct determines their raw score, which is then converted by a standard process into a scale score. This kind of scoring system is often called "objective" because it does not rely on human judgment to allocate marks.

By contrast, people have to make a judgment against a standard. In classrooms teachers are expected to combine evidence from a number of sources and make an overall teacher judgment. As we know from any situation where several people make an assessment of the same task, views of how the performance measures up will vary (think of TV judging panels). Moderation, including use of good exemplar materials, is a process that should help teachers become more consistent in their judgments. But shared understandings will take time to develop and the experience of the NCEA suggests no-one should under-estimate the professional learning involved.

In the script scrutiny exercise, a panel of experienced advisers and teachers with subject expertise in mathematics were given a set of PAT scripts at each of Years 4 to 8. They were tasked with going through the scripts and working out how the results would map in relation to the standards at each of the year levels. As they examined the scripts they analysed what questions each individual child got right and what they got wrong and made judgements based on the performance in the test as to where each individual was in relation to the standard. From this it became possible to indicate the range of scores at each year level that the judges felt were indicative of performance in mathematics well below, below, at, and above the National Standard.

It is possible the panel as a whole were overly tough. They may for instance, have placed too much weight on individual questions rather than looking at the overall script score. That would result in a student who had done well in the PAT test, but who had missed one or two key questions, being judged to fall short of the standard. On their own, individual answers to test questions are unreliable. It could be that on the day, a student unexpectedly got these questions incorrect, when in fact they have a strong understanding of the underlying concepts. The most powerful piece of information in terms of using a test such as the PAT as one of the sources of evidence for an overall teacher judgment is the overall score. The overall result combines information from all the questions in the test and does not lend itself to over-analysis of the performance on one or two questions.

It is our understanding that the script scrutiny work is to be revisited. Until this has happened we cannot say if the assessors were too tough, but nor can we rule it out as a possible explanation for the gap between normal progress on PAT and the progress expected by the National Standards for mathematics.

3. Could the level of the standard be set too high?

If the normal progress students make in mathematics really does fail to meet the National Standards, and we develop increasing confidence that the judgments that show this are accurate and consistent, the next question to ask is whether what we might now be expecting of our children is unrealistic. Is the standard too high to begin with? This question cannot be easily dismissed. It opens up a complex set of interconnected issues which we now explore.

Reflecting on the purpose of standards

We have been told that the performance levels described by the standards were developed by a process of working backwards from the performance expectations for NCEA Level 2. The idea is to keep as many of our children as possible, and preferably all of them, on an achievement trajectory that would see them able to succeed at NCEA Level 2. This would enable them to leave school able to go on to further learning pathways that open up a wider range of personal opportunities and enable them to contribute to the national prosperity that can come from having an educated and dynamic workforce. This seems an admirable aspiration but we have some questions about how it is enacted in practice.

One question thrown up by the equating gap between PAT: Mathematics and the National Standards in mathematics is whether the standards should signal appropriate overall academic progress or the kind of progress necessary for a specific *mathematics* trajectory. Put another way, does the progress we know that "average" students currently make (as measured by PAT) provide a satisfactory foundation for secondary school programmes leading to NCEA Level 2 success, regardless of whether students continue to take mathematics once it is no longer compulsory? Or, are we now saying that we expect most of our students to pass NCEA Level 2 mathematics courses, enabling them to go on to Level 3 and quite possibly to university mathematics?

These two alternatives surface important questions about the purpose of mathematics in the school curriculum, and they also lead to different sorts of questions about the "problem" for which these national standards might be argued to be the solution.

If we really do think that mathematics is so important that we should aspire to have every student pass a full NCEA Level 2 mathematics course (say 14 credits gained from relevant achievement standards) then the equating gap could signal a very real issue. It is not easy to work directly backwards from actual success in NCEA Level 2 mathematics to year 8 achievement levels to check for correspondence. Too few students currently actually take NCEA Level 2 mathematics courses for a general population check of correlations. However, we can see how actual results in NCEA Level 1 mathematics (when most students do take some form of the subject, with the majority focusing on numeracy-related standards) relate to NCEA Level 2 mathematics success rates. The calculations involved led to the broad estimate that somewhere between 1 in 5 and 1 in 7 of our young people could potentially pass a full NCEA Level 2 Mathematics course, if they chose to take the subject at this level. This range of between 14 to 20 percent is even lower than the equating exercise, which as noted, currently suggests that at Year 8 only

around 35 percent of students are actually reaching the National Standard (with the caveat about whether the equating exercise judgments were too tough, as outlined above).

Should we be concerned about this? Should we be telling many parents their children are now achieving below the standard when they have been making satisfactory progress, as this has been determined over the years, and as the PAT shows is "normal" for their age? If we really do think every child should be able to succeed in NCEA Level 2 Mathematics specifically, then the answer is "yes". But NCEA Level 2 Mathematics is not needed to access meaningful post-school education options, or essential to contribute to a dynamic and productive economy. We do need a good number of students who have high-level mathematics skills and thinking, but we do not need every student to have this level.

To reiterate one question we have already posed—should the standards set for the top end of primary school signal *general* readiness for future learning—a good academic foundation for later qualification success—or *specific* readiness to keep succeeding in mathematics? If we believe the former is the more important signal, then we can ask whether we have evidence that current norms of progress (as measured nationally by PAT and asTTle, and internationally by TIMSS) are a good foundation for secondary school learning in general. Clearly the answer is "yes" because we do not see wholesale failure of our middle-of-the-road students in secondary school. Currently, 71 percent of students gain NCEA Level 2: double the proportion the equating exercise judged would reach the National Mathematics Standard at Year 8

There will always be some middle-of-the-road students we would like to see push harder to achieve their potential but by and large these students do make good progress through secondary school.

There are important reasons to be concerned about our underachieving students but that is not the issue here—it is our average students' progress that has been called into question by the equating exercise.

Deeper issues are raised by the questions just posed. What is the purpose of mathematics in the secondary school curriculum? Traditionally it has served as a "gatekeeper" subject, used to sort out our most able students from those destined for different types of career pathways. Do we believe this situation has changed? If not, then it is clearly inconsistent to expect a subject to simultaneously sort to a standard and yet expect everyone to reach it. If we think things have changed and maths is no longer a "gatekeeper" subject (and these are good reasons for answering this way) then we need a much wider conversation about what it is. What is the place and purpose of mathematics in the curriculum and what are the implications for what maths is taught and how it is assessed? It is these kinds of conversations we need to have as a country looking to do things differently.