A new era for PAT Mathematics

—*Online, interactive, and adaptive*

ALEX NEILL

PAT: Mathematics tests have been around since 1974, and are used in thousands of New Zealand classrooms. In this article we outline the history of the tests, then describe some recent and ongoing work that marks a new era in the evolution of the assessments. Until now, PAT Maths tests were composed solely of multiple-choice items, and could only be administered, marked and analysed manually. The new era has already made tests available online, opening up a new world of possibilities for test items. Very soon we will introduce increasing numbers of interactive items and computer adaptive test administration algorithms.

The Progressive Achievement Test of Mathematics (PAT: Mathematics) has been developed specifically for use in New Zealand schools. Is is designed to help classroom teachers make informed decisions about their students' achievement and growth in the knowledge, skills, and understandings of mathematics and statistics as described in *The New Zealand Curriculum* (Ministry of Education, 2007) (*NZC*). Before 1970, New Zealand teachers had no such tests that were targeted specifically at their New Zealand students. The PATs have been through several revisions and updates over the past four decades.

A brief history of PAT: Mathematics

An initial pool of questions was developed in the early 1970s after wide consultation with a range of educators. These items underwent extensive testing in schools. A large-scale standardisation exercise was conducted on large sample of students. This ensured that the items met the criteria of being both valid and reliable (see Darr, 2005a and 2005b). The method of analysis used was the "classical test theory" because of the way it measures the reliability of the test. The tests provided normative information on student performance in the form of percentiles.

http://dx.doi.org/10.18296/set.0019

Since their introduction, the original PAT: Mathematics tests have been revised twice and new additional tests have been added to the pool. The first update occurred in 1993 and followed a similar standardisation protocol to the 1970s, with many of the same items being used as well as some new ones being developed. The concept of stanines was introduced to overcome some of the shortcomings of percentiles (though the latter were also included). Both the 1974 and 1993 series used two equivalent forms of the tests, each of which had 200 items.

The second revision, undertaken in 2006, was far more extensive. A small number of existing items were used, but the large majority were new, reflecting current mathematical thinking, and based on the developing structure of *NZC*. Seven tests were developed, and these were targeted from Year 4 through to Year 10.

The 2006 edition initially included one test per year level. Two additional tests were released soon after. Six new tests were released in 2015. This now gives schools the greater flexibility of having more than one test that is specifically targeted to each of Years 4 to 10.

The 2006 revision also gave schools the ability to directly measure student's mathematical progress over time. An alternative analysis was used[1] that allows results from all the test forms to be reported on the same measurement scale (measured in patm units). The scale measures students' mathematical ability directly rather than using normative information such as stanines and percentiles (though these are still provided).[2] Reporting all results on the same scale emphasis growth and allows progress to be tracked over time, as patm is an equal-interval scale. For example, a student who increases from 20 to 25 patm is considered to have made as much growth as one who moves from 50 to 55 patm.

Since 2006, schools have been able to use the New Zealand Council for Educational Research (NZCER) marking service and its online database of test results. These allow teachers and students to receive more immediate and tailored formative feedback. Schools can now choose to directly input their students' results into a website: www.nzcer.org.nz/tests/nzcer-marking. Alternatively, schools can send their student answer sheets to NZCER to have the information scanned and entered for them. In both cases, schools have access to an increasingly powerful range of reports that can assist them in interpreting their results in a variety of ways. These changes are shown summarised in Table 1, along with more recent and upcoming developments.

	Form of analysis and reporting	**Marking**
First release 1974	Classical (percentiles)	Hand marked
Updated 1993	Classical (percentiles and stanines)	Hand marked
New edition 2006	Rasch model (linear measurement scale)	Hand marked
Automated marking 2006	Rasch model (linear measurement scale)	Automated
Online administration 2008	Rasch model (linear measurement scale)	Automated
Two test per year level 2015	Rasch model (linear measurement scale)	Automated
Interactive and adaptive 2016	Rasch model (linear measurement scale)	Automated

TABLE 1. SUMMARY OF PAT DEVELOPMENTS

A new era of developments

Advances in information technology paired with students' increased access to the internet have enabled a new era for the PAT: Mathematics. It has opened up the potential for students to experience more active and personalised approaches to standardised assessment. Three specific developments are tests that can be online, interactive, and adaptive.

Online

A recent innovation allows schools to administer the tests in an online environment. Currently, any of the published PATs can be administered online on a wide range of devices. This has the distinct advantage that marking is done automatically, and with 100 percent accuracy. Studies have shown that hand marking is an error-prone process as well as being time consuming (see for example Chapman, St. George, & Ibell, 1985). The full range of reports are also available automatically. NZCER's internal monitoring of data collected online and using paper-and-pencil versions of the test of PAT: Mathematics tests has shown that the mode of administration of multiple choice items (either pencil-and-paper or online) has little effect on the difficulty of individual items and tests (Darr, 2014).

Interactive

Moving to an online platform is allowing NZCER to develop a wider range of question types beyond the multiple-choice questions currently used. Initially, four new item types are being developed. Each of these will be automatically marked by the computer. The new question

types are described below, and examples of them are shown in Figure 1.

1. Image hotspot (click on the correct number of responses). Students click on all the correct responses.
2. Numerical responses (fill-in-the-box). Here, the student inputs their own answer to a question into a box rather than selecting a given response. The computer ascertains if the response is correct.
3. Image grouping (drag and drop). This allows a student to drag some images onto another image in a correct way.
4. Sorting. The student drags the icons into the correct order. The icons automatically snap into place.

Over time, more and more new types of items will be added to the pool of PAT: Mathematics items. These will only be available online and will be a feature of the adaptive environment that is described in the next section. We anticipate that these questions will prove to be engaging for the student, and will give the opportunity for richer information on student responses to be available, especially from the numerical response questions.

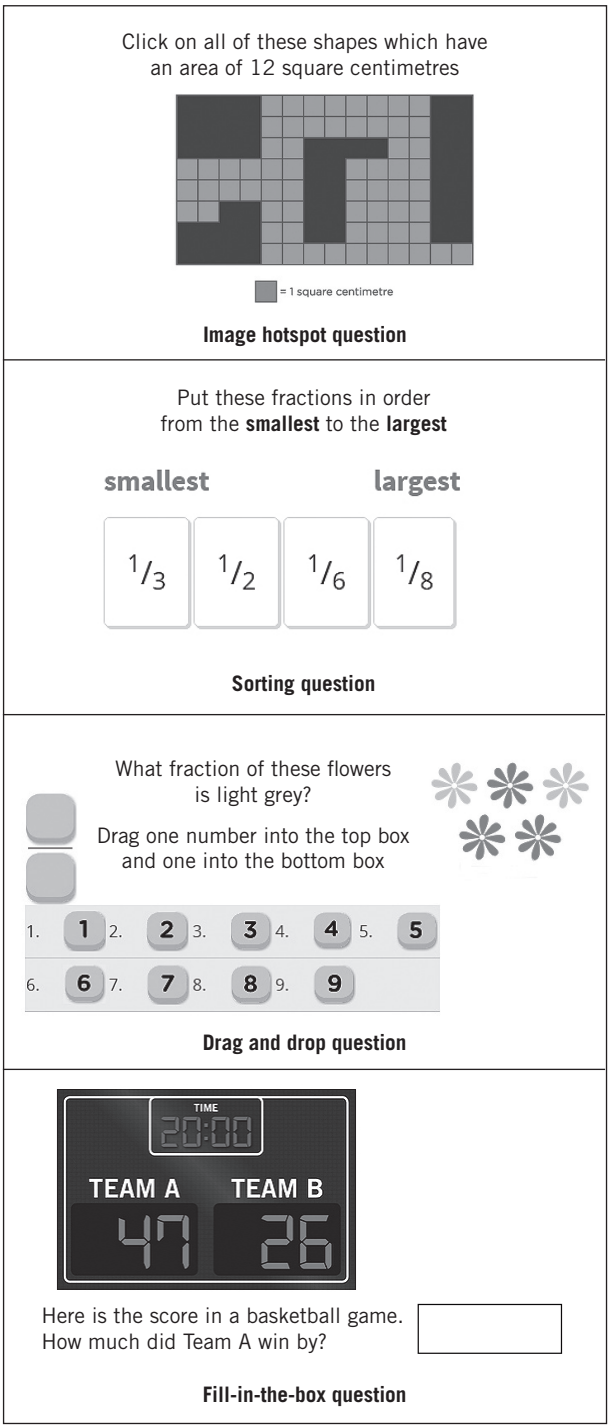

FIGURE 1. NEW ITEM TYPES

Adaptive

Computer adaptive tests allow students to be administered a test which is individualised for them. The student is given an initial question. If the student is successful with this, the computer selects a more difficult question. If the student gets it wrong, an easier question is selected. This process continues a number of times until the student is being given questions at about his or her level of ability. This can be seen diagrammatically in Figure 2. It has the advantage that the test will be well targeted, providing more precise information about how the student is tracking and ensuring the test is not too difficult or too easy.

The adaptive tests will require new forms of reporting. Some existing reports will still be valid, particularly at the individual student level. NZCER is current exploring other forms of reporting that will be specifically tailored at adaptive tests, and will allow teachers to look at trends within and between classes.

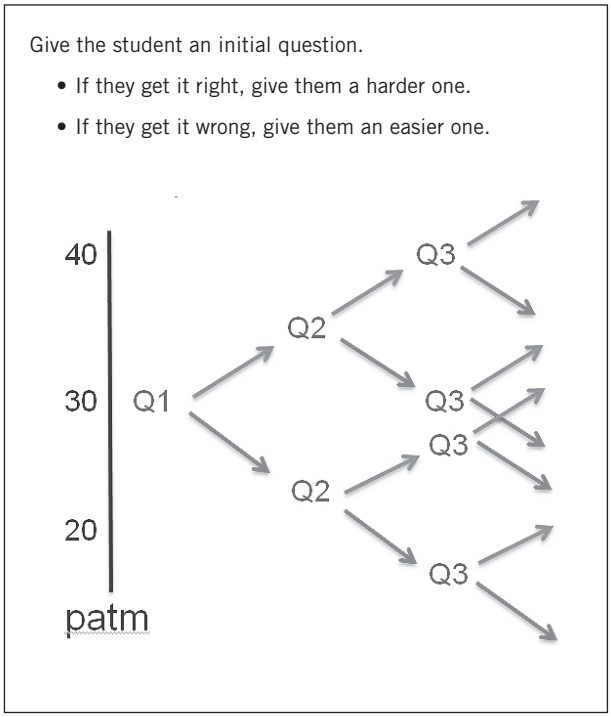

FIGURE 2. SCHEMATIC REPRESENTATION OF AN ADAPTIVE TEST

Concluding comments

Over the 40 years that NZCER has been providing PATs to schools, the tests have evolved, especially over the last decade. They were initially an exclusively "by-hand" approach to administration, marking, and reporting. The affordances of modern technology are allowing students to engage with the tests online, and experience a wider range of question types. Teachers and schools are also able to explore the results in far greater depth. It has been highly rewarding, as well as challenging, being involved with the changes made since the mid-2000s. The team at NZCER have valued the way that so many of your schools have been willing to assist us in developing these tests, and we look forward to continued development and ongoing contact with you.

More information

To obtain more information on PATs or to order them, visit the NZCER website: www.nzcer.org.nz/tests/pats

Notes

1 The underpinning psychometric model changed from classical test theory to a model built on item response theory—the Rasch model. Both of these address test reliability from differing viewpoints.

2 Initially, percentiles were the only method of reporting. Percentiles, however, have issues around their correct interpretation, and were prone to be used to overstate fine differences in achievement levels. Stanines are more robust to this, and schools have largely moved towards using them. Stanines can, however, only be reported relative to the time of year the test was given, and the year level of the student. NZCER now strongly recommends that schools move their thinking towards using the patm measurement scale. This is invariant to the time of testing, or to the year level of the student.

References

Chapman, J. W., St. George, R., & Ibell, R. (1985). Error rate differences in teacher marking of the Progressive Achievement Tests: Sex, ability, SES and ethnicity effects. *New Zealand Journal of Educational Studies*, *20*(2), 165–169.

Darr, C. (2005a). A hitchhiker's guide to validity. *set: Research Information for Teachers, 2*, 55–56.

Darr, C. (2005b). A hitchhiker's guide to reliability. *set: Research Information for Teachers*, *3*, 59–60.

Darr, C. (2014). Computer-administered vs pencil-and-paper tests: Is there a difference? *set: Research Information for Teachers*, *3*, 61–64.

Ministry of Education. (2007). *The New Zealand curriculum*. Wellington: Learning Media.

Alex Neill is a researcher and resource developer at the New Zealand Council for Educational Research.

Email: Alex.Neill@nzcer.org.nz