The transformational change of adaptive assessment

Dr Jarrod Hingston and Toby Newton

24 August 2022
Embarking on change

Technology is driving opportunities to advance how we assess and measure achievement, ability and aptitude. Moving to computer-based assessment has provided a whole new set of processes and infrastructure that has enabled ACER to deliver more effective assessment products and services such as the online versions of the Progressive Achievement Test, more commonly referred to as PAT. More recently change has occurred through the development of a computer adaptive assessment system within ACER’s existing Online Assessment and Reporting Platform. Today we are reflecting on these advancements from an organisational change perspective.

According to the Harvard Business School, Organizational change is the action a business takes to change any of its underlying components, such as processes, culture, people, product, infrastructure, or technology. In the case of ACER’s development of a new computer adaptive assessment system, most if not all of the underlying components underwent some change. It was the magnitude and pace of change that went beyond expectations.

(Harvard Business School Online (2020), Common Types Of Organizational Change & How To Manage Them, URL: https://online.hbs.edu/blog/post/types-of-
organizational-change#:~:text=Very%20often%20transformational%20change%20refers%2C%20is%20often%20sudden%20and%20dramatic, accessed on 30 June 2022).
In order to reflect on the project from an organisational change perspective, we need to understand the anticipated degree of organisational change and the actual degree of organisational change.

Harvard Business School described adaptive changes as small, incremental adjustments that organizations and managers make to adapt to daily, weekly, and monthly business challenges. These changes are often related to fine-tuning existing processes, products, and company culture.

When the project was first being conceptualised in 2017, there was a broad expectation that the project could be accomplished through a fine-tuning of existing products, processes and technology.

There were existing examples of ACER projects that had used an adaptive change approach to introduce adaptive assessment into products and services. The fine-tuning required software development changes on existing systems that would hard code adaptive pathways into test sittings.
Problems with adaptive change

Prior to the project initiation, through ACER’s Centre for Assessment, Reporting and Innovation, some substantial research and effort was invested in by ACER in looking at different models of adaptive assessment and what made them effective.

This included making a series of recommendations for how ACER should best advance assessment products and services using adaptive assessment.

In trying to plan for an adaptive change or fine-tuning approach a multitude of shortcomings were identified that would prevent the project from meeting the recommendations of the research.
Problems with adaptive change

One of the shortcomings identified was reporting.

The pre-existing reports were all developed to provide teachers with easy to access information to efficiently identify strengths and weaknesses in students’ performance. But these reports were developed assuming that students would be sitting the same assessment and attempting to respond to the same test items.

Applying the fine-tuning model to these reports quickly revealed that data would present across a multitude of pages and screens and would no longer offer teachers the same benefits.

Other shortcomings of fine-tuning included limiting opportunities for improvement such as automated test level allocation and being able to draw on new or parallel content to limit students’ exposure to the same test items.

Consultation with schools, systems and experts identified that this needed to be a bigger change.
Harvard Business School described Transformational change, on the other end of the spectrum to adaptive change.

Transformational changes are typically much grander in scope than incremental, adaptive changes. Transformational change may even refer to a dramatic evolution of some basic structure.

Shifting our thinking from Adaptive Change to Transformational Change enabled us to consider the project as building a new and innovative adaptive assessment system as opposed to an adaptive assessment that would need on-going incremental changes. It enabled us to identify opportunities for improvement, consult with schools and seriously consider their advice and build a service based on research recommendations.
Taking a transformative change approach led to some significant changes in the scope of the project. A working group that drew on expertise from across ACER was formed to oversee the planning and execution of the project.

After reviewing the project requirements four distinct workstreams were created being Branching, Content, Test Assign and Scoring and Reporting. A further workstream known as user experience was created that would cut across all of the other workstreams.
The branching workstream entailed a range of work packages. The first piece of work involved testing recommendations from the previous research into the preferred branching model. Existing data were modelled to identify which of the recommendations would most likely lead to optimal results for a PAT Adaptive assessment in Reading and Mathematics. The models supported a three stage tailored test design that would provide a good range of items and content balance to be presented across three testlets, with branching occurring at the end of the first and second testlets.

The second work package in this workstream was to create a new system within ACER’s Online Assessment and Reporting Platform that would support the delivery of test content in testlets and enable branching at the end of a testlet based on lookup tables that could be changed if required. This was an important part of the transformational change aspect as previously branching rules were hard coded and could not easily be tweaked or modified. A proof of concept trial was conducted in September 2019 to demonstrate that this work package had been successful. This was a key gateway for the project to continue as per the transformational plan.

A further work package was also delivered that would enable parallel testlets to be introduced to the system. This included the Adaptive assessment system identifying if
students had previously undertaken a testlet and if so, the system would then direct them to a new testlet of similar difficulty and content balance. The parallel testlet functionality would also support the creation of testlets and pathways for special accessibility requirements.
The Content workstream involved creating blueprints for the PAT Maths Adaptive and PAT Reading Adaptive assessments, including reviewing content appropriateness, test length and strand balance for various achievement levels. Test developers conducted item inventories against the blueprint, identifying what types of Mathematics and Reading items needed to be developed and trialled.

New test items were developed and trialled in Australian schools.
One of the major transformative change aims of the Adaptive assessment system was to automate the assignment of the assessment’s entry level. The branching design of the PAT Adaptive assessments earmarked eight entry levels and the intention was to develop a functionality that would select a student’s entry level based on a previous PAT scale score or their year level.

This process would run at the same time as a teacher assigned a PAT Adaptive test to the student. The test assign workstream focused on developing this functionality and also providing mechanisms to override the automated entry level.
The largest and most challenging of the workstreams was the reporting workstream. An adaptive change approach to fine-tune pre-existing reporting was quickly identified as not going to be suitable for the Adaptive assessment system. It was also identified that a transformational change approach would open opportunities to innovate and improve the reporting of students’ results. Through consultations with teachers, schools and experts the following priorities were identified for reporting:

- Reports needed to be interactive and interconnected
- The PAT scale needed to be clearly displayed in charts and figures
- Detailed student reports and the ability to view test items were valued by teachers
- Data needed to be exportable.

In a few moments Toby will demonstrate how these priorities were delivered and the reporting work became a system within a system.
A user experience workstream was established that cut across each of the other workstreams. The aim of user experience is to focus on making the adaptive assessment system clear and easy to use. This required consideration of all aspects on the platform from the user interface to processes required to report on an assessment. Much of this work was led by Toby and led to some valuable innovation.
Jarrod has explained the vision we had when we set out to undertake a transformational change in our online assessments.

And the developments we have made as a result will continue to be an ongoing endeavour, so I'm not spoiling the ending by briefly jumping forward to this year.

With PAT Adaptive assessments having been in use by schools for just 12 months, ACER was named a Finalist for the 2022 International eAssessment Awards in the category of 'Best Transformational Project'.

This was a huge honour and recognised the value of the work and the positive effects PAT Adaptive has already had on students, teachers, and ACER as an organisation. It also validated our decision not simply to iterate on an existing product but to aim towards a truly transformational change in the way we approach assessment and reporting.

We don't have time to discuss in detail each of the workstreams that Jarrod just mentioned, so I'd like to focus particularly on how we approached reporting in this new adaptive assessment context.
As Jarrod alluded to, we recognised early that computer adaptive testing would not be adequately served by an adaptive change to our existing online reports. A transformational change was required.

- The nature of computer adaptive assessments, with students presented widely varied and personalised sets of items, meant a simple summary of correct and incorrect responses is impossible.

- Our analytics showed at the time teachers were relying on this existing summary-style report at a rate 8 or 9x higher than any other report available. It was by de facto the PAT report.

- As a result, many teachers were missing out on the full range and depth of information available from PAT assessments.
A new approach to reporting

While the innovations of our new computer adaptive assessment model would be largely invisible to teachers, the reporting would be where its value is seen and felt directly.

So we had to completely re-think our approach, in the first instance to address the immediate challenges of a new assessment structure, and more broadly to give teachers the fullest access to all aspects of the data.

Just as the existing PAT described achievement scales provided us a way to automate the accurate assigning of test levels to students, the PAT scales would form the centrepoint of all of our new online reports – not only emphasising their utility in expressing and correlating both student achievement and the difficulty of test items, but also providing an anchor for a consistent visual design that would improve user experience and navigation between different views of the data.

Aware of the huge discrepancy in use between some of the existing PAT reports, our new approach was to move away from separate, standalone reports to a single, integrated data suite that would place teachers directly in their results data and allow them to explore as needed.
From the very beginning, we had some guiding principles intended to place user experience at the centre of the new design:

• Teachers should not be required to make a series of blind decisions – such as dates, student groups, tests, or report types – before being presented with the data.

• Rather than feeling like discrete components, with separate entrances and exits, the reports should simply function as different ways of approaching or slicing the same set of data.

• The visual appearance and layout of each report should be consistent and familiar, including the peripheral buttons, links and interactive elements used to navigate or change the data parameters.

• The naming of each report should be clear and communicate the nature of the information and its use, but should not be relied upon to aid navigation.
From the early design phase we went into development, during which, we undertook two separate periods of user testing and feedback with schools from around Australia to ensure both the new assessments and the reporting elements were meeting the needs of teachers and students. This was crucial to identify areas of improvement in both the technical design as well as the user experience.

The result is the PAT Data Explorer. But rather than a finished product, we are deliberately considering it an ongoing project that continues to be added to and expanded.
**Student:** Farrell, Ray  
**Current year level:** Year 7  
**Progress:** 2018: 121.5  
**2019:** 127.0  
**2020:** 124.5  
**2021:** 124.0
This is the Item Performance report, which is our solution to the problem of analysing group performance across a wide range of items in an adaptive assessment.

You can see that it takes advantage of the PAT scales for each learning area to assist teachers in analysing the data. Each icon represents an item seen by at least one student in the selected group.

While some items will have been seen by more students than others, and amongst a group the total number of items seen may be in the hundreds, it's possible to view and interact with them effectively in a way that would not be possible in a traditional table of students vs items.

Displaying the items visually in this way, makes it easier to parse and interpret the information at a glance. The relative difficulty of the items is indicated by their vertical position in the chart (higher items are more difficult than lower items), whilst the rate of correct responses is communicated by the icon colour and up/down arrows. Typically, you would expect the more difficult items to have a lower correct response rate and vice versa, so teachers can quickly identify anomalies that may
warrant closer attention. For example, relatively easy items with low correct response rates.

Importantly, the filters at the left allow teachers to focus on students not just by year level or class group, but by overall achievement, which is not something that's possible in our older reports.

In the case of the same item being seen by two students of varying ability, it may not be surprising or particularly useful to see that one answered correctly and the other didn't. But if you correct for this by focusing just on students whose overall ability is similar, than the discrepancies and anomalies in their responses become much more meaningful.
Item: Recognises that multiplying a decimal fraction by 100 moves the decimal point two places to the right

<table>
<thead>
<tr>
<th>Strand</th>
<th>Sub-strand</th>
<th>Proficiency</th>
<th>Difficulty</th>
<th>Correct responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number and algebra</td>
<td>Fractions and decimals</td>
<td>Fluency</td>
<td>131.5</td>
<td>15</td>
</tr>
</tbody>
</table>

Scale score: 131.5
Band: 13
All: 15
Year level: 5
Tags: 
Date: 
PAT Reading resources

Group by

- Band

Currently grouped by Australian Curriculum codes
- Resources are targeted at band levels

- 150 and above
- 140-149
- 130-139
- 120-129

- 110-119

Text complexity example 110

Teaching activities

Classifying and re-organising
- Grouping similar information into a table

Navigating text structures
- Using adjectives (extension)

Rewording
- Describing numbers (extension)
- Matching clauses

Scanning for key information
- Computing information in a procedural text (extension)
- Scanning a narrative

Using grammatical structures
- Personal pronouns (extension)
- Teaching personal pronouns (extension)
Internally, the development of the PAT Adaptive assessments and the PAT Data Explorer has provided ACER with the technical capacity to develop and deliver any number of new computer adaptive assessments, because the structure was deliberately built to be expandable and configurable, not just something retrofit to an existing assessment.

The development process also afforded ACER the opportunity to audit and re-think our approach to item development and management. As a result, we now have the capacity to compile psychometrically robust tests that meet a variety of content requirements, including gender and cultural representation, web accessibility, and other audience needs.

The capacity to automatically assign appropriate tests to students based on ability estimates has the potential to benefit future assessments, both linear and computer adaptive. And the PAT Data Explorer is already being expanded to support additional reporting features and other online assessments.

More broadly, as a result of our experience with PAT Adaptive, we are now incorporating explicit user research efforts as a fundamental element of our ongoing work.
The ultimate aim of assessment is to uncover where students are in their learning progress – to identify what they can do and what they cannot yet do, so teaching and learning can be shaped and targeted effectively.

So for students, although the inner workings of the computer adaptive model are largely invisible, PAT Adaptive ensures that they are presented test content that is well-targeted to their abilities.

Every student is appropriately challenged, so their areas of need can be identified. But just as importantly, every student has an opportunity to demonstrate what they can do.

This means, greater levels of engagement and a more accurate picture of every student's level of achievement at the time of the test.

Beyond this, parallel testlets that ensure students are not exposed to repeated content in consecutive tests, and accessible test pathways that support the needs of students with visual impairments and other disabilities, expand the flexibility and accessibility of the assessments.
For teachers, the changes brought by PAT Adaptive mean it is now less work to implement truly effective assessment of student learning.

ACER has always championed the importance of meeting students at their point of need. In a traditional assessment context, this requires professional judgement and, most importantly, time to choose and assign appropriate test levels to every student. It is far easier to instead assign all students in a class or year level the same test, which may be suitable for the 'middle', but will likely disengage both the top and bottom achievers and return less-than-useful data.

The automatic test assign element of the PAT Adaptive work has completely removed this barrier to good assessment practice. Teachers can spend more time on the things that matter, like effectively using the results data to improve student learning.
Last year, Term 4 – typically our busiest period – saw a total of 1.4M 'old' reports generated by teachers using PAT. The simple, item–student summary, 'Group' report accounted for 83% of this total. During the same period, 422K PAT Data Explorer reports were generated.

In Term 2 2022, our quietest period for testing and reporting, 381K 'old' reports were generated. Again, 85% of these were the 'Group' report. By contrast, over half a million PAT Data Explorer reports were generated.

More than just speaking to the success of the PAT Adaptive project, this increase in usage of the PAT Data Explorer, especially outside of our peak periods, most importantly means that teachers are being exposed to – and using – a greater variety of report types to analyse and explore their data from different angles.
Unsurprisingly, the most heavily used reports are the 'Achievement' report, the landing page for the PAT Data Explorer, which summarises students' overall achievement against the PAT scale, and the individual Student reports, which allow for close diagnostic analysis of every student's performance. Together, these account for approximately 60% of reports generated.

Given the power of the PAT Adaptive assessment model to accurately target every student and elicit quality information about their current level of achievement, it is heartening to see the extent to which teachers are focusing on the Student reports.

But there is also significant use of each of the other report types, particularly those that provide insight into student performance at the item level, diagnostic information that can support targeted planning. The Band report is also well used, which suggests that we are succeeding in our aim of helping schools better understand and use the PAT scales and described achievement bands to interpret student achievement and plan next steps.

In terms of the contribution we hoped PAT Adaptive would have in schools, our goal was to have 1000 schools take up PAT Adaptive in its first 12 months. Within just nine months, more than 3000 Australian schools were using PAT Adaptive and the PAT
Data Explorer, including 1500 Victorian government schools, through a state-wide contract with the Department of Education.

This level of take-up and the enthusiasm teachers and school leaders have shown for PAT Adaptive has been extraordinary, and speaks to the transformational change achieved by PAT Adaptive.
Thank you

Dr Jarrod Hingston and Toby Newton

24 August 2022