

**Education and Skills Committee
Scottish National Standardised Assessments**

Dr Craig Jones- New Zealand

Tumuaki Tuarua/ Deputy Secretary, Te Puna Mōhioitanga/ Evidence, Data & Knowledge

Examples of low-stakes national standardised tests. Eg, when are they administered (what age/school year & when in the year), how are they administered, and what are the purposes of the collecting the information?

How is the information used?

Are these cohort or sample based tests?

International assessment studies (TIMSS, PIRLS, PISA): We assume you are familiar with the purpose and details of implementation of these assessments, but they are one of the most important datasets we generate for national monitoring purposes.

National Monitoring Study of Student Assessment: This is the successor to the National Education Monitoring Project, which was referred to in the OECD literature review of standardised tests you referenced. Its focus is to determine students' achievement against the New Zealand Curriculum (English Medium). A representative sample of students in Year 4 (approx. age 9) and Year 8 (approx. age 13) are assessed in two different learning areas each year, meaning each learning area gets assessed approximately every six years. The purpose is entirely for national reporting – only about 25 students from each school are selected to participate, and so there is generally not enough data generated to inform teaching or school planning. For more information about NMSSA and the types of reports generated, see: <http://nmssa.otago.ac.nz/about-us.htm>

e-asTTle: This is a tool that can be used to generate a range of standardised assessments across Years 4 (approx. age 9) to 10 (approx. age 15). Teachers select specific topics within reading, writing and mathematics (such as algebra or geometry) and curriculum level, and the e-asTTle tool will provide students with questions tailored to that topic and level. Use of e-asTTle by schools, kura and teachers is optional and free of charge. It can be used across whole classes or in a targeted fashion for individual students. It is used by approximately one-third of primary schools, and slightly less secondary schools. E-asTTle is primarily intended to provide formative information for teachers. However, the tool asks for consent to send data generated by the tool to the Ministry of Education for research purposes (and most users provide this consent). Our initial analysis showed that data from this collection is broadly representative and is consistent with representative assessments (such as NMSSA), so we are beginning to use this data more heavily at the national level for research purposes

(see <https://www.educationcounts.govt.nz/publications/series/he-whakaaro/understanding-student-attainment-and-progress> for an example of this).

For more information about e-asTTle, see: <http://e-asttle.tki.org.nz>

Observation Survey of Early Literacy Achievement (Third Edition)/Burt Word Reading Test: These are NZ developed assessments of early reading that are used in a large number of schools. They are not compulsory to use. The Observation Survey is generally used after one year of school instruction (generally age six in New Zealand), although it is normed for four age groups between the ages of five and seven. It is primarily used to determine a student's grasp of basic reading and writing concepts and skills, and to pinpoint any gaps in understanding. The Burt Word reading test is a standardised test of reading before eight or nine years. It was normed for use in NZ in 1981 and places attainment within an age band. The Ministry receives these scores for the students who participate in Reading Recovery (the most common reading support programme), but this is a highly unrepresentative sample. For examples of the national reporting for the Reading Recovery sample, see: <https://www.educationcounts.govt.nz/statistics/schooling/reading-recovery>. For more on the Observation Survey, see <http://assessment.tki.org.nz/Assessment-tools-resources/Commonly-used-assessments/Observation-survey>

There are also a large range of other standardised tests (the most popular of which are the PAT and STAR) that are used by schools to inform both teaching and school planning/evaluations. However, this information is not collected centrally and so we do not consider it "national". Further information about these tests can be found at <https://www.nzcer.org.nz/tests/school-tests>

Are the tests used for more than one purpose? What are the strengths and drawbacks of this?

These assessments tend to be used mainly for the purpose for which they were created. This is primarily due to practicalities. For example, TIMSS/PIRLS/PISA/NMSSA are of limited use for formative assessment because only a small number of students in each school participate (and the results are in the form of plausible values, which are difficult to interpret for individual students).

E-asTTle, the observation study, and other assessments such as PAT used by schools are not representative samples (although e-asTTle comes close, and we are beginning to use this more) and/or we don't centrally collect the data for all students. This means that they cannot be used for some functions. For example, the Observation Survey is useful for determining who receives Reading Recovery and monitoring progress through Reading Recovery, but cannot be used for evaluation purposes or shed insight into the reading ability of all New Zealand students.

Because assessments have tended to be designed for defined purposes (either national reporting/research or formative assessment), this has the strength that they are easier to create and may be more valid. However, it also has the limitation that it increases the burden of data collection. For example, the lack of an assessment covering all students means that most evaluation activity needs to contain a collection of primary data, as opposed to relying on existing administrative collections. This increases the costs of valuable activities such as research and programme evaluation.

Are there tensions between collecting robust, reliable and valid data and supporting learning? How are these resolved?

Definitely. New Zealand has a decentralised schooling system, with each school governed by its own community. This means that there is a lot of resistance to any data collection perceived as mandatory and/or “one-size-fits-all” (e.g. a previous “National Standards” measure for reading writing and Maths which have recently been removed). Any data collection that attempts to address these concerns by being opt-in or flexible in terms of its assessment design runs into the risks of selection bias or validity issues.

This is a topical issue in New Zealand at the moment, with these and related tensions currently being considered by an independent Ministerial Advisory Group (MAG). The MAG is due to report back to the Minister on how the school sector can better collect and use data by mid-2019. For more about the work of the MAG, see: <https://conversation.education.govt.nz/conversations/curriculum-progress-and-achievement>

Is there evidence that low-stakes national standardised tests improve outcomes?

We feel that there is strong evidence of the benefit of using accurate and valid formative assessment information to inform teacher practice, as well as for school planning and evaluative purposes. We assume you are aware of this literature, but can provide you with some citations if that would help. Because many of our assessments are opt-in, we take their continued use to be a signal that schools and teachers consider them to be valuable.

In terms of the sample-based national monitoring studies, their primary benefit is in shedding light on particular problem areas that may otherwise go unnoticed. For example, NMSSA, e-asTTle, and TIMSS data all point towards a drop-off in progress relative to curriculum expectations in the middle schooling years. As a result of this data, this period is a particular focus for our policy. However, because these sample-based studies are meaningful only at the national level, it is difficult to isolate the benefits that they provide to individual students.