# EXPLORING THE IMPACT OF A LARGE-SCALE DIAGNOSTIC SCIENCE TEST AND FORMATIVE PRACTICES. A mixed-methods study.

James Scott MEd

Doctor of Philosophy C02041

University of Technology Sydney

Faculty of Arts and Social Sciences

**Certificate of original authorship**

I, James Scott declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy by Thesis in the Faculty of Arts and Social Sciences at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

Production Note:
Signature removed prior to publication.

29 August 2018

**Acknowledgments**

This thesis would not have happened without insights, support and encouragement from a number of people.

My thanks to Dagmar McCloughan, ESSA Team Leader, for the opportunity to be involved with the ESSA program in the early years of its development and implementation. Also, my thanks to Professor John Pegg (University of New England) and Associate Professor Debra Panizzon (then from the University of New England) for the opportunity to be a part of the research team investigating the potential of SOLO as a tool for improving assessment for learning. After that initial involvement, both provided me with their advice, encouragement and support which I sought at different times whilst I worked on this thesis.

I would like to acknowledge Doctor Geoff Barnes who encouraged me to run with the idea that the residual from a regression procedure was a measure of a real effect of teaching. My thanks also to the NSW Department of Education for providing me with access to ESSA and NAPLAN data in a form that I could use for the purposes of this thesis. Particular thanks in this regard are due to Doctor Nadine Smith and former colleague and dear friend Gerry McCloughan.

Associate Professors Nick Hopwood and Tapan Rai from the University of Technology Sydney have my gratitude for the time and advice they provided as I developed the approach I wanted to take with the research model and analysis of data.

I am extremely grateful to the science teachers who responded to the survey about their practice and particularly so to the teachers who made themselves available to participate in the case studies. I have undertaken to provide them with the results of my work in a form that I hope will be useful to them.

I had the support and advice of two excellent supervisors, both at the University of Technology Sydney, for this thesis. Professor Peter Aubusson who encouraged me to undertake this project in the first instance and Associate Professor Matthew

Kearney who took over in the later stages to assist me bring it to a conclusion. My gratitude and thanks to both for their patience, advice and support.

This thesis also had the benefit of the considerable editing skills of Doctor Terry Fitzgerald who is also at the University of Technology Sydney.

Finally, I want to acknowledge the forbearance of Daune my wife who, in the end, waited patiently and supportively for me to complete this project so that we could resume our lives together.

**Thesis format**

This is a conventional thesis comprised of title, front matter, glossaries (acronyms and terms used), table of contents, list of figures, list of tables, abstract, six chapters, appendices and references consulted in the preparation of this thesis.

## List of Acronyms

| | |
|---|---|
| AAS | Australian Academy of Science |
| ABS | Australian Bureau of Statistics |
| ACARA | Australian Curriculum Assessment and Reporting Authority |
| ACCI | Australian Chamber of Commerce and Industry |
| ACER | Australian Council for Educational Research |
| AE | At Expectation (see also WAE and WBE) |
| ANOVA | Analysis of Variance |
| AQF | Australian Qualifications Framework |
| ARG | Assessment Reform Group |
| BCA | Business Council of Australia |
| BOS | Board of Studies |
| BOSTES | Board of Studies, Teaching and Educational Standards |
| CC | Curriculum Corporation |
| CCII | Centre for Continuous Instructional Improvement |
| DEC | NSW Department of Education and Communities |
| DET | NSW Department of Education and Training |
| D of E | Department of Education |
| ESA | Education Services Australia |
| ESSA | Essential Secondary Science Assessment |
| EV | Acronym for the acronyms ESSA and VALID. |
| F | The Foundation or entry level for schooling (see K). |
| HSC | Higher School Certificate |
| ICSEA | Index of Community Socio-Educational Advantage |
| K | Kindergarten or entry level for schooling (see F). |
| NAP-SL | National Assessment Plan-Scientific Literacy |
| NAPLAN | National Assessment Plan Literacy And Numeracy |
| NESA | New South Wales Education Standards Authority |
| NGSS | Next Generation Science Standards (US) |

| | |
|---|---|
| NSES | National Science Education Standards (US) |
| OECD | Organisation for Economic Co-operation and Development |
| PCK | Pedagogical Content Knowledge |
| PIRLS | Progress in International Reading Literacy Study |
| PISA | Programme for International Student Assessment |
| SEA | Socio-Educational Advantage |
| SEAR | Science Education Assessment Resource |
| SET | Science, Engineering and Technology |
| SLPM | Scientific Literacy Progress Map |
| SMART | Schools Measurement Assessment and Reporting Toolkit |
| SME | Science, Mathematics and Engineering |
| SOLO | Structure of the Observed Learning Outcome |
| SPSS | Statistical Package for the Social Sciences |
| STEM | Science, Technology, Engineering and Mathematics |
| TIMSS | Trends In Mathematics and Science Study |
| US | United States of America |
| VALID | Validation of Assessment for Learning and Individual Development |
| VET | Vocational Education and Training |
| WAE | Well Above Expectation (see also AE and WBE) |
| WBE | Well Below Expectation (see also AE and WAE) |

**Glossary of terms as used in this thesis**

| | |
|---|---|
| artifact | Something made by human effort, in this context related to educational assessment. |
| assessment as learning | Assessment as learning occurs when students are their own assessors. Students monitor their own learning, ask questions and use a range of strategies to decide what they know and can do, and how to use assessment for new learning. (NESA, 2018) |
| assessment for learning | Assessment for learning involves teachers using evidence about students' knowledge, understanding and skills to inform their teaching. Sometimes referred to as 'formative assessment', it usually occurs throughout the teaching and learning process to clarify student learning and understanding. (NESA, 2018) |
| assessment of learning | The use of evidence of learning to make a summative judgment of achievement against outcomes and standards. Sometimes referred to as 'summative assessment'. It usually occurs after a period of instruction. The judgment is often expressed as a mark, percentage or grade. The usefulness of the grade or mark depends on validity and reliability of the processes used to gather and assign value to the evidence gathered. (NESA, 2018) |
| assessment-related work | Is the purposeful collecting of evidence of learning, creating the means by which that evidence was obtained (if not by direct observation of behaviour), the assumptions used to interpret that evidence, the choice of text forms used to represent and communicate results of assessment, and subsequent uses for those results. |

| | |
|---|---|
| capabilities | A measure of the ability, capacity, power or potential to do something. The *Australian Curriculum, Science* includes seven general capabilities all students are expected to acquire as they progress through schooling. |
| Curriculum Corporation | A national educational support entity created by the Federal, state and territory governments in Australia to produce educational resources for Australian Schools. It was replaced by Education Services Australia (ESA) from 2010. |
| competencies | See capabilities. |
| curriculum | The documents teachers use to inform the learning activities they plan and deliver to students. |
| diagnostic assessment | Gathering evidence of learning to identify gaps, strengths and weaknesses in student learning. |
| education jurisdiction | States and territories in Australia manage the delivery of educational services to students in Australia. They provide for registration and regulation of public and private schools in their geographic areas of jurisdiction. |
| educational standards | Are the learning goals students are expected to achieve, usually after set periods of instruction typically associated with Year or Grade levels. |
| feedback | Information provided by an agent regarding aspects of one's performance or understanding. |
| formative assessment | See assessment for learning. |
| formative practices | Instruction informed by formative feedback. |
| high stakes assessment | Any assessment where the results have consequences for the recipient of those results. |
| key competencies | A set of competencies related to equipping students for work. |

| | |
|---|---|
| low stakes assessment | The use of evidence of learning in ways that reduces to a minimum unintended, usually negative, consequences for the learner. |
| outcomes | Measurable or observable behaviours intended as a result of instruction. |
| Primary Connections | A set of curriculum materials produced by the Australian Academy of Science designed to assist K-6 teachers to teach science. |
| proficiency areas | Areas of skill or expertise. |
| proficiency levels | Descriptions of response features that differentiate between levels of skill or expertise. |
| regression | Regression is a statistical process for estimating the relationships between variables. |
| Science by Doing | A curriculum support resource produced for secondary science teachers by the Australian Academy of Science. |
| scientific literacy | Scientific literacy is the ability to engage with science-related issues, and with the ideas of science, as a reflective citizen (OECD). It is also the specialized literacies that distinguish science literacy from general literacy and numeracy. |
| SEA quarters | Socio-Educational Advantage (SEA) proportions, relative to Australia, in school populations. (ACARA *MySchool* website) |
| SEA score | Socio-Educational Advantage (SEA) score is a composite measure of socio-educational advantage generated for the purposes of this project. |
| selective entry schools | A category of school in NSW, entry to which is determined by student results in tests of reading, mathematics, general ability and writing. |

| | |
|---|---|
| self-regulated learners | Students who can plan their own learning, monitor their performance and then reflect on the outcome of that learning. |
| Skills, cognitive | Include remembering, thinking logically and reasoning, explaining and describing. |
| Skills, employability | Skills related to communicating, working in teams, problem solving, initiative and enterprise, planning and organising and self-management. |
| Skills, generic | Groups of skills variously described as basic/fundamental, people-related, conceptual/thinking, personal skills and attributes, skills related to the business world and skills related to the community. |
| SOLO model | Structure of the Observed Learning Outcome (SOLO) theory that involves two learning cycles within a mode of thinking |
| SOLO taxonomy | Structure of the Observed Learning Outcome (SOLO) theory that describes a single learning cycle within a mode of thinking |
| standards framework | Descriptions of levels of performance in a number of categories relating to curriculum, teaching or other profession. |
| statistically significant | Is the probability of finding a given deviation from a null hypothesis, or a more extreme one, in a sample. (SPSS definition) |
| STEM system | Science, Technology, Engineering and Mathematics institutions in a country or larger group that prepares people for work in, and including, the institutions that produce STEM outputs in society and related economies. |
| summative assessment | See assessment of learning. |

| | |
|---|---|
| syllabus | A detailed curriculum that in NSW may be used to define the scope of an external test. |
| The Board | A generic term for the statutory authority in NSW with responsibility for determining the curriculum and related assessment requirements schools need to comply with so that students satisfy requirements for receipt of credentials. In the course of this project that authority began as the NSW Board of Studies (BOS), became the NSW Board of Studies Teaching and Educational Standards (BOSTES) before becoming the NSW Education Standards Authority (NESA) in 2017. |
| The Department | A generic term covering the NSW government authority responsible for delivering public education services to students in NSW. It went from being at the beginning of this project (2012) the NSW Department of Education and Training (DET) to the Department of Education and Communities (DEC) to the NSW Department of Education (D of E). |
| Year 8 | The year of schooling in Australia (Grade in other places); in this case the ninth year of schooling. |

## Table of contents

## List of Figures

## List of Tables

**Abstract**

Researchers working with schools in the UK and elsewhere are finding that explicitly teaching students the "five strategies of formative assessment" (Black and Wiliam, 2009, p. 8) is helping to re-engage students with science. This thesis presents findings about the impact of two major interventions on the assessment-related work of junior secondary science teachers in the New South Wales government school system (the largest in Australia) and on student science results. The first intervention took the form of advice to teachers about formative assessment in the official science curriculum (introduced in 2003), where it is called assessment for learning. The second took the form of a mandatory low-stakes, large-scale, test-based diagnostic assessment program involving Year 8 students. This program was fully implemented across NSW from 2007. The assessment framework used to inform the development of test items and tasks and that informs the comprehensive feedback provided to students, parents and teachers is underpinned by Structure of the Observed Learning Outcome (SOLO) theory. Three research questions guided data collection. The research design employed mixed methods, including both quantitative and qualitative methods as well as case studies involving sixteen purposively chosen school sites. Descriptive and inferential statistics were applied to the analysis of both state-wide and school-specific, teacher-provided survey data about their practices and school-level test results. An interpretive approach was used to generate assessment-related work narratives from audio-recorded interviews and artefacts of assessment practice provided to the researcher by volunteering science teachers in the case study schools. The findings show that teacher use of three of five dimensions of formative practice and an explicit focus on teaching students the skills of writing to learn science produced science test results that were above expectation. Less certain was the hoped-for finding that students were also acquiring the skills of learning how to learn. An unexpected finding was that students in regional schools where science results were well above expectation were less positive about their school science experience than their metropolitan counterparts.