Mixed Methods Research to Investigate the Impact of Formative Assessment on Science Teaching in NSW Schools

Presentation at CAR seminar 7-8 November 2018

by Jim Scott, PhD candidate

Deakin Downtown, 727 Collins St., Melbourne

OVERVIEW

- Context
- Three research questions
- Design…mixed methods in three phases / all three phases had quantitative analysis & findings / third phase included qualitative interpretation and inferences
- Data collected from Department of Education records / *MySchool* website / School level data / teacher interviews and teacher selected assessment-related artifacts
- Findings
- Selected bibliography (see printout)

Context

- Assessment of…for…as learning / classroom assessment / formative assessment / feedback / learning how to learn
- Science syllabus goal in NSW – produce independent learners (BOS, 2003)
- Two initiatives are the focus of research in this project
    1. Assessment for learning was promoted to teachers in NSW syllabuses, including science
    2. Science was added to literacy and numeracy testing (Year 8 from 2007). The primary purpose of these tests was diagnostic.
- Five strategies of formative assessment underpin ***the theory of formative practice***…which is the theoretical framework used to explore the impact of the two initiatives on science teaching in NSW
- The assessment framework for the test included both syllabus standards and Structure of the Observed Learning Outcome (SOLO) theory

Assessment for learning, formative practices and self-regulated learning

ASSESSMENT FOR LEARNING	DIMENSIONS OF FORMATIVE PRACTICE
1. is an essential and integrated part of teaching and learning 2. reflects a belief that all students can improve 3. involves setting learning goals with students 4. helps students know and recognise the standards they are aiming for 5. involves students in self-assessment and peer-assessment 6. provides feedback that helps students understand the next steps in learning and plan how to achieve them 7. involves teachers, students and parents in reflecting on assessment data. (BOS, 2003, p. 70)	1. clarifying and sharing science learning intentions and success criteria (LISC) 2. engineering effective science classroom discourse and using learning tasks that elicit evidence of student learning (CDEL) 3. providing feedback that moves learners forward in their learning of science (FTAL) 4. activating students as instructional resources in science for one another (and the teacher), including peer assessment (ASIR) 5. activating students (and teachers) as the owners of their own science learning, including self-assessment (ASTL).

The ESSA / VALID test

- Was pen and paper in 2007; online from 2010 (first in Aus to do so)
- Purpose is to provide feedback on learning progress
- Progress in terms of curriculum (syllabus) standards, defined in terms of outcomes and related content, and six levels of thinking described by the Structure of the Learning Outcome (SOLO) model
- It has extended response and short response items (the test 'looks and feels' to students like NAPLAN and other tests)
- It includes a student survey to find out what students think about the test, about science and their experience of science at school
- Students and parents receive an individual report of progress in March April the following year
- Comprehensive feedback to schools at the level of individual student responses to each task and item in a software package that has a limited capacity to manipulate results
- Science results are retained within the school (unlike literacy and numeracy results which are now published on the *MySchool* website).

Research questions

- Motivating question for me: what impact have the two initiatives, promoting assessment for learning and a diagnostic science test, had on the assessment-related work of junior secondary science teachers in NSW government schools?

- Three research questions…
  1. What use are science teachers making of the test-based assessment program including SOLO and why is it used or not used?
  2. What formative practices are evident in the work of science teachers and why are they used or not used?
  3. Is the use of formative practices by teachers linked to improvement in students' Year 8 science results and later achievement in and engagement with science?

The research design…

- Mixed methods after Cresswell and Plano Clark (2011)

    1. convergent parallel design
    2. explanatory sequential design
    3. exploratory sequential design
    4. embedded design
    5. transformative design
    6. multiphase design.

- Transformative design chosen using explanatory sequential design with transformative intent

1st phase: sample selection

- Four years of science data (2011-2014)
- Regression of science over NAPLAN predictor (394 out of 465 schools with 10 or more Year 8 students)
- Sort schools based on residual from 1 to 394
- Top, middle and bottom 20% of schools (WAE, AE and WBE science results)
- All WAE, AE and WBE schools invited to complete an anonymous online survey
- At the end, asked to provide and EOI to participate in a case study.

Thinking behind 1st phase

- Contributions to “accounted for variation” in student test results (Hattie, 2003) / student background (50%) / school factors (20%) / teacher (30%)
- **Scientific literacy attainment = EV test results – general literacy and numeracy skills contribution**
- NAPLAN-result predictor / coefficient of determination $R^2_{av} = 0.892$ / residual is measure of scientific literacy attainment & ‘effect’ of teaching
- Independent measure of student factors (*MySchool* socio-educational disadvantage / advantage profiles of case study schools)
- Nothing in the data to indicate why the residuals were as they were.

2nd phase: survey analysis

- Small and unequal numbers in samples: WBE (32), AE (28) and WAE (25)
- SPSS software used…parametric and non-parametric options
- Group response data tested for normality and homogeneity of variance
- Whole group then categories within groups (such as five dimensions of formative practice for Q9 – 15)
- Mostly used non-parametric ANOVA (Kruskal-Wallace test) and Games-Howell multiple comparisons tests to confirm means differences that were significant
- Descriptive statistics when group means were not significantly different.

Survey structure

A. About accessing and using the components of the test
B. About SOLO
C. About assessment for learning (five dimensions of formative practice)
D. About you and the school

- Invitation to express an EOI for case study (need to self-identify)
- 258 invitations to schools / 101 survey returns / 85 usable /42 teachers & 36 schools self-identified

Table 4.17

Sample items from the online survey with a teacher or student focus

Teacher focus	Student focus
Q9c explain to students the indicators or success criteria I will be looking for in their work	Q9d allow students some input in deciding what success criteria are to be applied
Q10h I explain my responses / thinking	Q9f ask students why they think they are being asked to do the proposed activities
Q10f I use test or assignment items and tasks as stimulus for discussion (in class)	Q9g encourage peer feedback based on success criteria
Q11e (provide feedback) advice about how to improve	Q10d ask students to explain their thinking
Q12c (feedback) refers to misconceptions	Q10e use the "think-pair-share-report" strategy
Q14c I evaluate lessons and record ideas for change next time	Q13d (students) self-assess by redoing work to a higher standard
Q14f, g & h access and use information in class…about assessment for learning	Q13e (student self-) selection of items for a portfolio
Q15a collaborate with my science teacher colleagues to develop a shared understanding of what progression in science learning looks like	Q13f self-assess by getting them to keep a journal of their reflections in their own words (on what they have learned in science lessons)
	Q14a students give feedback on my teaching

3rd phase: case study

- Case study schools of “maximum variation cases” Flyvbjerg (2011) to provide good examples for professional learning.
- Third phase case studies involving quantitative and qualitative data
  - ✓School results—EV, Year 10 and Year 12 senior science course completions
  - ✓Assessment artifacts—models of ‘best’ practice (teacher selected)
  - ✓Semi-structured Interviews—based on school & science department assessment practices and priorities, formative practice dimensions and the online survey (triangulation)
  - ✓Assessment-related work narratives (interview questions provided the scaffold for interpretive analysis
  - ✓Correlations to test strength of associations.

Research question 3

Is the use of formative practices by teachers linked to improvement in students' Year 8 science results and later achievement in and engagement with science?

What does engagement and achievement look like in comparable schools with very different residuals?

Is there any evidence that a greater proportion of students in schools more frequently exposed to formative practices were autonomous, self-regulated learners?

ENGAGEMENT… six items from the student survey

A. I want to study a science subject in Years 11 and 12.
B. Science is the hardest subject I learn.
C. In primary school, I enjoyed lessons that were about science.
D. In secondary school, I enjoy science lessons.
E. Students asked to nominate their three favourite subjects (15 were listed including science).
F. Students asked to nominate the three subjects they thought they learned most in (15 options listed including science).

Three predictions…

Research question three: Is the use of formative practices by teachers linked to improvement in students' EV results and later achievement in and engagement with science?

Predictions

1. Overall Year 8 science test results for students in comparable schools will be better in WAE schools than AE schools, and AE school results will be better than WBE schools.
2. Overall Year 10 science result patterns for students in comparable schools will be better in WAE schools than AE schools, and AE school result patterns will be better than WBE school patterns.
3. The proportion (relative to English) of students completing Year 12 science courses in comparable schools will be highest in WAE schools, and AE schools will have a higher proportion of completions than WBE schools.

Findings relating to research questions one and two

- There were statistically significant differences between WBE and WAE teachers in the frequencies of their use of the resources related to the science test and in their understanding of the purpose of the science test
- There were no statistically significant differences between the three groups of teachers when it came to understanding of and overall engagement with SOLO (very low scores)
- There were statistically significant differences between WBE and WAE teachers in the frequencies of their use of three of the five dimensions of formative practice
  - ✓ Discourse that elicits evidence of learning
  - ✓ feedback to students that advances learning
  - ✓ Their use of and modelling to colleagues and students of good learning behaviours (cf PEEL) including self-assessment
- There were no significant differences between the groups in relation to student ownership of dimensions of formative practice

(Some of the) findings in relation to research question three

For case study schools overall:

- No findings could be made in relation to self-regulation of students in any of the three school groups
- WAE schools all had strong science literacy programs running in their science departments in the early secondary years
- At the end of Year 8 top band students in WAE provincial case study schools had the lowest satisfaction scores for their school experience of science (compared to metropolitan and selective schools)…but they had very good science completion scores at the end of Year 12
- At the end of Year 8 the fully selective case study schools rated science as their hardest subject.