Article

Teaching, learning, and assessment of science investigation in Year 11: Teachers' response to NCEA

Azra Moeed[1]* and Cedric Hall[2]
[1]School of Educational Policy and Implementation and [2]School of Educational Psychology and Pedagogy, Faculty of Education, Victoria University of Wellington, PO Box 17 310, Wellington

In New Zealand, the school science curriculum promotes open-ended science investigation, but the assessment regime in year 11 requires investigation with 'direction' and a more limited understanding of investigation. This case study research explored teaching practice and teacher understanding of science investigation, and the connectedness between learning, motivation to learn, and assessment. This paper presents the findings on teacher practice related to the teaching and assessment of science investigation. The results indicate that teacher practice of science investigation changed in response to the internal assessment requirements associated with science investigation for the National Certificate of Educational Achievement (NCEA). The nature of this change raises validity and reliability issues for the assessment of student learning of science investigation.

Introduction

Learning in scientific investigation is an important goal of science education, alongside the acquisition of scientific knowledge, understanding, and practice (Kanari & Millar 2004). The practical aspect of the subject has had a distinct and central role in science curricula internationally. Science educators have argued that there are benefits in engaging students in practical activities in science (Abrahams 2011; Hofstein 2004; Hofstein *et al*. 2008; Lunetta 1998; Tytler 2007; Woolnough 1991; Wellington 2005) and suggest that by carrying out an investigation students learn the related science concepts and understand better the nature of science (Hodson 1990; Roberts & Gott 2006).

Woolnough (1991) argued that scientific investigation promotes a holistic approach to learning science through the linking of scientific concepts to the process and outcomes of an investigation. In the view of Patrick & Yoon (2004), students gain most from science investigation when they 'discuss expectations, observations, conclusions, theories, and explanations before, during, and after conducting the activity' (p. 319). Millar (2004), in agreement with this view, argued that student learning in relation to science investigation needs to be seen as a recursive process rather than a constrained procedure. The notion of a recursive process was also clearly promoted in *Science in the New Zealand Curriculum* (Ministry of Education 1993):

> *The processes of investigation are not sequential. The process may begin at any point ... will tend to move backwards and forwards. Students should be reflecting on their decisions, actions, and findings and modifying their plans and actions as they are proceeding.* (p. 47)

In New Zealand, Achievement Standard Science AS1.1 defines investigation as:

> *…an activity covering the complete process: planning, collecting and processing data, interpreting, and reporting on the investigation. It will involve the student in the collection of primary data.* (New Zealand Qualifications Authority 2005, p. 3)

Investigation, as mandated by the New Zealand science curriculum, promotes a recursive process where the students consider and modify their plans as the investigation proceeds. However, the assessment process related to AS1.1 effectively requires a linear and sequential process because of its focus on one 'fair testing' type of investigation in which students control a single variable. In a New Zealand case study of science investigation, Hume & Coll (2008) concluded that students in year 11 were acquiring a narrow view of science investigation as fair testing, and that although learning was taking place, students' responses demonstrated rote learning and low-level thinking. A related concern was raised by Allen (2008), who argued that an investigation needs to be challenging for the student. If the

* *Correspondence: azra.moeed@vuw.ac.nz*

Azra Moeed is a senior lecturer and curriculum leader of science education in the Faculty of Education at Victoria University of Wellington. Her teaching and research interests include science teacher education and teaching and learning of science in New Zealand schools.

Cedric Hall is an Emeritus Professor of Education at Victoria University of Wellington. His teaching and research interests cover curriculum, teaching, learning, assessment, evaluation, and the design of professional development and learning. He has published several papers relating to assessment in New Zealand, including critiques of NCEA.

activity is too simple or the answer is known in advance, there is little intellectual challenge left in the activity. The student then focuses on getting the right answer rather than carrying out a scientific investigation.

In a British study, Cleaves & Toplis (2007) researched students' views of learning and assessment of science investigation and reported that students considered that their teachers trained them to do investigation. Students said that the teachers told them 'this is what you have to do' (p. 92) and 'this is what you need to write' (p. 92) to get a good mark. Additionally, students said they were taught to repeat data collection; they knew that they had to do this but did not understand why. Students also said they learnt to comment about anomalous results; they considered that it was good to have anomalous results because if they could explain them they would gain a better grade. Cleaves & Toplis (2007) reported that students develop a view that investigation is a part of science that they have to learn in order to get marks, rather than a view that science is 'predicated upon investigation' (p. 92).

Assessment theory

The two main guiding principles that underpin assessment are validity and reliability (Hall 2007). As noted by Hall (2007, p. 6):

> *Validity focuses on 'fitness for purpose'. A valid assessment task is one that fulfils its intended purpose(s), such as fairly testing the course objectives and content, fostering student learning, and motivating further interest in the subject.*

However, within the context of school science at the senior secondary level there are at least three key elements that need to be considered in judging the validity of the assessment of science investigation. Firstly, the intent and requirements of the New Zealand Science Curriculum need to be understood and interpreted appropriately by everyone involved in teaching and assessing science investigation. Secondly, the translation of the curriculum into the relevant National Certificate in Educational Achievement (NCEA) standards (AS1.1 for year 11) requires close scrutiny to see how well the standards maintain the integrity of the curriculum. Thirdly, the actual assessment tasks that are undertaken by students need to match well the intent of the relevant assessment standards and the curriculum from which the standards are derived. The literature reviewed in this section suggests that there are question marks surrounding all three elements. The present research looks at assessment practice for AS1.1 to uncover more fully what problems exist and the implications of these for future teaching and assessment practice.

In addition to validity, assessments need to be reliable. As noted by Hall (2007, p. 7):

> *Whereas validity focuses on 'fitness for purpose', reliability refers to the extent to which the assessment provides an accurate measurement of each student's understanding or learning'.*

The point to note is that an assessment may appear to be valid because the material being tested focuses on important knowledge and skills. However, the results may not be reliable because the measurements fail to capture students' learning accurately, as may happen if tasks contain ambiguities, or markers interpret and apply criteria differently, or students behave inconsistently (e.g. misinterpret a question or are influenced by factors related to health or emotional state). In general, reliability is commonly evaluated through evidence of the consistency and stability of results across assessors and over time (Harlen 2005). In relation to science investigations, students' performance can differ from one investigation to another depending on the content, complexity and openness of task, the processes used to oberve and record students' performances, the consistency of markers in applying criteria, and factors related to the context of the assessment such as the teaching approach taken by teachers.

The relationship between validity and reliability is well demonstrated by Roberts & Gott (2003) commenting on Sc1, a similar assessment requirement in the UK to Achievement Standard AS1.1. They suggested that a validity issue arises where an assessment involves the observation by a single teacher of a large number of students carrying out a practical investigation. They noted that complex tasks could take two to four hours to perform, typically resulting in few investigations being carried out:

> *Sc1 has become routine, with a limited number of cases assessed. In some instances, Sc1 coursework has become so formulaic that performance is more akin to the recall of a complex protocol than the creative solution of a problem.* (p. 104)

In other words what is being assessed is not actually what was intended to be assessed. In Gott & Duggan's (2002) view, to get a valid indication of a student's ability to carry out an investigation, results over a number of assessments (covering different types of tasks) would need to be combined. The point to note is that such a procedure might also improve the reliability of the results because it is based on several assessments of science investigation, enabling a more stable judgment to be obtained.

An important factor noted, or at least implied, in some of the literature reviewed in this section, is that the assessment of science investigation, because of its 'high stakes' nature, influences the teaching approach adopted in schools (this is often referred to as the 'backwash' effect of assessment). The risk that exists is that teachers 'train' their students to pass the assessment, thereby losing focus of the important understandings that students need to engage with (such as the relationship of science investigation to science knowledge more widely and the role of problem-solving in discovering or applying science knowledge). There is no denying that assessment has an important place in teaching and learning in school science. What is debatable is whether teaching to the test or task is a desirable outcome of an assessment system. The research of Cleaves & Toplis (2007) in the UK and Hume & Colls (2009) in New Zealand indicates that this approach is commonly adopted.

The New Zealand Science Curriculum (Ministry of Education 1993) promotes open-ended science investigation, where students have control over defining a problem, choosing the method, and arriving at solutions (Simon *et al*. 1992), but the assessment regime in year 11 requires investigation with direction, which is likely to be associated with a more limited understanding of investigation.

This paper reports research that studied teachers' response to this contradictory situation by investigating how year 11 science teachers practise science investigation. It was part of a larger project that looked at the phenomenon of science investigation focusing on the key elements of motivation to learn, learning, and assessment (Moeed 2010).

Methodology

The research reported here adopted case study methodology, drawing on qualitative data to study the phenomenon of 'science investigation'. The intention was to understand investigation through those who practise it in their unique contexts and through the interactions that take place in that setting (Merriam 1998). The research involved a critique of the *National Curriculum Framework* and *Science in the New Zealand Curriculum*, a regional survey of year 11 science teachers, interviews with year 11 science teachers, and an in-depth study of one science class (see Figure 1).

Teacher survey

An anonymous postal survey was administered to year 11 science teachers in the Wellington region; 101 teachers completed the survey, representing a 61% response rate. Participants in the survey were drawn from coeducational schools (64%), boys' schools (17%), and girls' schools (19%). The teaching experience of participants varied – 40% had fewer than five years' experience, 14% between 6 and 10 years, 17% between 11 and 15 years, and 29% had 16 years or more. One respondent did not indicate teaching experience. Forty-six percent of participants were drawn from high socio-economic communities (deciles 8–10), 33% from middle socio-economic communities (deciles 4–7), and 21% from low socio-economic communities (deciles 1–3). Fifty percent of participants came from large schools (student numbers 800+), 32% from medium-sized schools (500–799 students), and 19% from small schools (fewer than 500 students). Sixty-one teachers were female and 40 male.

The main survey themes relevant to the data reported here focused on:

- teacher practice of teaching science investigation
- change in teaching of science investigation since internal assessment for NCEA
- reasons for any change in practice
- preparing students for assessment including formative assessment
- the procedure for assessment of science investigation.

Teacher interviews

All ten teachers who taught a year 11 science class in a typical coeducational, medium-sized school situated in a middle socio-economic community were interviewed through a semi-structured interview. The main question themes were:

- teachers' views about teaching and learning science investigation
- teachers' approach to assessment
- change in the teaching of science investigation since internal assessment for NCEA
- the approach for assessment of science investigation for AS1.1.

Full details of the postal survey and the teacher interviews are provided in Moeed (2010).

Results

The regional survey results, and study school science teacher interviews showed that year 11 science teachers focused on training their students to undertake the fair testing type of investigation in preparation for internal assessment of science investigation. The approaches the regional teachers said they used included 'repetition', 'doing tasks similar to those assessed' and 'practising fair testing'. This approach was also adopted by the study school science teachers, who said they were 'training' their students to investigate and 'getting them to go through the hoops'. Some of these teachers reported an emphasis on students learning the skills needed to investigate. Thus procedural knowledge rather than procedural understanding and conceptual learning were deemed appropriate preparation for AS1.1.

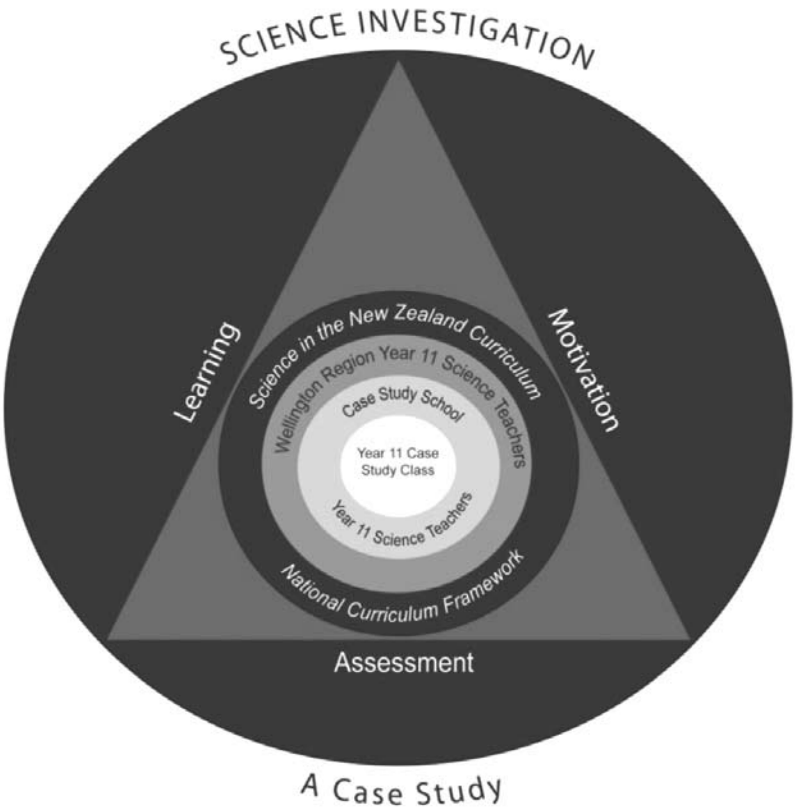

Figure 1. The case study of science investigation in year 11 science.

Science teachers in the study school said that the teaching approach they took was contrary to how they would ideally teach science investigation but in the interest of students' achievement. Teaching 'what would be assessed was seen as a pragmatic solution to the dilemma they faced – there was little choice given the assessment regime in place. The training approach to teaching investigation was reinforced by teachers constantly using the template designed for AS1.1.

Teacher survey

Teachers were asked to indicate if they had taught science before the introduction of the NCEA and whether their practice of teaching science investigation had changed since its introduction and, if so, to explain in what way. Sixty-six respondents (65%) had taught year 11 science before the introduction of NCEA. Fifty-five (83%) of these reported a change in their practice of teaching science investigation after its introduction. Eight of these 55 teachers reported doing more science investigations, 22 the same number, and 25 fewer.

Change in practice

Teachers provided multiple ways in which their practice had changed in relation to investigation and assessment. Investigation-related changes made up 66% of the responses. These included teachers saying that they did more complete investigations (31% of responses), and that the investigation had become compartmentalised (13% of responses). Twenty percent of the responses were assessment-related, a typical response being:

Investigations become an exercise in fulfilling criteria for credits. (Teacher 036)

A further 11% of responses recorded that teachers did more holistic investigation. One teacher said:

The process of doing a complete investigation can take up to three lessons. Kids do the planning task in one lesson and then wait to do the investigation (gather data) in the next lesson and then I either get them to write the report for home work or it has to be done in the next lesson. Sometimes they forget what they had done the last time. To me this complete investigation feels like more compartmentalised than complete. (Teacher 073)

A small number expressed concern that teachers were doing fewer student-initiated investigations. Although this was a very small percentage (6%), it is important as open-ended investigation is student-initiated and usually based on something that the student wants to find out.

Some teachers indicated that their practice included teaching the students the language required to get a particular grade:

Emphasis on small things, in other words do these things and you will get an A [Achieved], M [Merit] or E [Excellence]. (Teacher 069)

Concern was also expressed that there was less time available to do other practical activities.

The reasons for change in practice were coded under the five categories: learning; assessment; less time; student motivation; and future use. Reasons coded as learning included responses such as ‘students need to learn to investigate in science because it is a practical subject’ (Teacher 016). Assessment reasons offered included ‘they need it because it will be assessed’ (Teacher 023) or ‘for achieving in AS1.1’ (Teacher 032 & Teacher 079). Teachers who said there was less time to do investigation gave reasons such as ‘each investigation takes several lessons to complete so there is less time for investigations’ (Teacher 082). Some teachers said they did investigation as ‘students like doing them’, ‘enjoy them’ or they are ‘more motivated when doing investigations’ (Teacher 093). A few teachers reported reasons such as ‘students need to know how to investigate for science in senior school’ (Teacher 056).

Teachers who had taught more investigations since the introduction of the NCEA were concerned with assessment and student learning, but also said they had less time than before. Those teachers who were doing the same number of investigations as they did before the NCEA had similar concerns about time; however, they stated that doing science investigation had motivational benefits (although these teachers did the same number of investigations, they said that how they taught investigation had changed). The teachers who were doing fewer science investigations were concerned about assessment and the lack of time but offered motivational reasons for not doing them (Table 1).

Table 1. Reasons given for change in practice in teaching science investigation after NCEA was introduced.*

Reason for change	Percentages for the responses by teachers who conducted, after NCEA was introduced:		
	More investigations	The same number of investigations	Fewer investigations
Learning	37	21	4
Assessment	42	41	50
Less time	21	18	40
Student motivation	0	15	6
Future use	0	5	0

* It should be noted that many teachers supplied more than one reason. The percentages are based on the number of responses made by each group, not the number of respondents.

Meeting assessment requirements was the most frequent reason given for change in the number of science investigations, whether the teachers were doing more, the same, or fewer than they had before the introduction of NCEA. The next factor was having less time. It is noteworthy that those doing more investigations considered enhancing student learning as the second most important reason. Learning becomes less important for those doing the same number of investigations but becomes still less when teachers choose to do fewer investigations.

More than a quarter of the responses (28%) indicated that teachers prepared their students for AS1.1 by doing tasks similar to those used for assessment and using the template from the Ministry of Education website, Te Kete Ipurangi (TKI).[1] Another quarter of the responses indicated that teachers used fair testing type tasks. Only 16% of responses recorded that teachers used formative assessment and gave students feedback as to how they could improve. Other responses indicated that they prepared their students by teaching them the skills of planning, interpreting and processing information, and reporting. Some indicated that they started preparing students from year 9 and familiarising them with the terminology used for AS1.1 (Table 2).

Separately from the data collected for Table 2, 78% of teachers carried out practice assessments in the form of a ‘mock examination’ or ‘trial run’. Their reasons for doing these assessments were to prepare students and familiarise them with the assessment context, enhance investigative skills, and provide feedback to improve their performance. Some teachers indicated that through such assessments they could identify and address students’ alternative conceptions. In their view, these assess-

[1] Te Kete Ipurangi (2005). Watch that car go. Retrieved 8 March 2010 from http://www.tki.org.nz/e/search/results.php?1%3Aelem=DC.Subject.Classification&1%3Aval=NCEA%3BNCEA%20Science&1%3Avalop=AND&1%3Asearchtype=term&2%3Aelem=TKI.Level&2%3Aval=NCEA+Level+1&2%3Avalop=AND&2%3Asearchtype=term&xsl_lang=en&xsl_path=/search/results_e.php

Table 2. Teachers' reported student preparation for AS1.1 (97 teachers made 189 responses).

Student preparation	Teacher responses	
	Number	Percentage
Doing tasks similar to those assessed	53	28
Practise fair testing	47	25
Formative assessment and giving feedback	30	16
By teaching skills needed for investigation	22	11
Start preparing students from year 9	18	10
Teach the science concepts	17	9
Do lots of practical work	2	1

ments increased student motivation and confidence. Since these trial tests represent a way of providing formative feedback to students, the relatively low figure for 'Formative assessment and giving feedback' in Table 2 (16%) should be seen as a response to a particular question about student preparation and not a figure about all forms of formative assessment.

Teachers saw disadvantages for themselves from carrying out formative assessments in terms of the workload associated with the preparation and marking (38% of responses). Some teachers (19% of responses) were concerned about the ethical issues in giving students too much help through formative assessment. Some saw assessment as demotivational if the task was too difficult.

Teacher interviews

When talking about the goals for student learning through investigation, all teachers (n = 10) focused on the fair testing type of investigation that is assessed in AS1.1. Overall, the goal was to teach a 'fair testing type of investigation' and foster the associated learning skills; however, concern was expressed that students were doing investigation for which they already knew the answer.

Teachers (n = 10) said that students learn the process skills of planning, gathering information, processing and interpreting information, and reporting findings, which are those identified in the assessment guide for AS1.1. Four teachers mentioned that they taught what students need to write in order to get an Achieved, Merit, or Excellence grade. Two teachers said they wanted the students to know that science is real and that we investigate all the time.

None of the teachers interviewed were satisfied with the process followed for the assessment of AS1.1 (following the requirement of fair testing, controlling variables and following steps to get to an answer already known). Their reasons were different but each expressed a genuine concern for their students which was obvious during the interview. They were 'despondent', 'upset', 'not impressed', 'uneasy', questioned the 'fairness', and 'pragmatic' – saying 'this assessment had to be done'.

All teachers said they taught students to investigate based on the requirements for AS1.1. All teachers gave students the opportunity to do at least one formative assessment (trial run), which was the school's science department's policy and was very similar to the task students were going to be assessed on.

One teacher explained that if the task for formative assessment was not sufficiently similar to the one used for AS1.1, most students would struggle and require a lot of help from the teacher. In his view, he was giving the students 'too much help' for AS1.1. However, as the school was now using the format required for AS1.1 in years 9 and 10 (a simplified AS1.1 template), he felt that this could change the practice of having to teach to the AS1.1 assessment.

Change in practice

Only half of the teachers interviewed had taught before the introduction of the NCEA and so only those teachers could comment on this aspect, and all five said that their practice had changed, by:

- focusing on fair testing type of investigation
- teaching students to use the template for the assessment of AS1.1. One teacher prepared students to write their answers in the right place in the template
- using a learning task almost identical to the assessed task
- using the procedure required for assessment
- putting all learning tasks in the NCEA format
- doing formative assessment as a trial run and providing feedback about what to write to improve the grades.

Raising her concern, one teacher said, 'the requirement to do and write in a particular way is leading to the template approach in writing up the investigation'. Another concern was that of the subject context in which the assessment was carried out. Teachers had taught students how to investigate when they were teaching the same subject as the one in which students would be assessed (e.g. in physics). However, when it came to the actual assessment for AS1.1, more than half the teachers (n = 7) were teaching another topic and students were disadvantaged because the assessment was out of context. One teacher said:

> *I was teaching chemistry at this point but this college runs AS1.1 in the exam week, so I was teaching chemistry and somebody else is teaching physics and somebody else is teaching biology and the context of the assessment is a physics one. We are assessing out of context. It is not fair to my students.*

All teachers said that once the assessment was over they did not give investigation the same amount of time (i.e. three lessons). Teachers (n = 8) said they asked students to investigate and just write the plan and results. Others (n = 2) said that they still insisted on fair testing and looking at the reliability of the data. Two teachers said that once the assessment was done, students were just not interested in doing practical work so they put the time into preparing for the examination. One commented:

> *To be perfectly honest we lead up a lot to that and then after that I'm still very insistent on reliability and fair testing, but things kind of start to flag a bit after that.*

Discussion

Investigation in practice: Fair testing

The regional survey results and study school science teacher interviews showed that year 11 science teachers focused on training their students to undertake the fair testing type of investigation in preparation for internal assessment of science investigation. The approaches used by the teachers in the regional survey included 'repetition', 'doing tasks similar to those assessed', and 'practising fair testing'. This approach was also taken by the study school science teachers. Some of these teachers reported an emphasis on students learning the skills needed to investigate. It would appear that procedural knowledge rather than procedural understanding and conceptual learning were deemed appropriate preparation for AS1.1.

Science teachers in the study school said that the approach they adopted to teaching science investigation was contrary to how they would ideally teach this curriculum area but in the interest of students' achievement and because students had to be assessed, teachers were pragmatic and continued to teach 'what would be assessed'. A view was that there was no choice. This training approach was reinforced by constantly using the template designed for AS1.1.

Regionally, and in the study school, more fair testing investigations were carried out when teaching physics or chemistry topics than biology or astronomy topics. According to Tytler (2007), such an imbalance occurs because it is easier to control variables in physics and chemistry. Evidence from this study suggests that in a fair testing investigation as practised in year 11, the design aspects of scientific investigation (planning) were reduced to the notion of variable control, where the student was making a comparison between two options and controlling variables to test a hypothesis. In the view of both Lunetta *et al.* (2007) and Tytler (2007), investigating in mainly physics and chemistry contexts is problematic, as potentially it could lead to students thinking that investigation is only done in these subjects.

A particularly influential factor for the focus of science teachers on fair testing is that the assessed investigation for NCEA Level 1 is a fair testing type of investigation. Although other types of investigation, including pattern seeking, classifying and exploration, are included in the curriculum, they are not specifically assessed in NCEA, which raises the issue that if other types of investigation are not formally assessed, they are less likely to be taught. More importantly, if students mostly experience fair testing they are likely to have a limited view of science investigation (Hume & Coll 2008).

Some teachers in the study school said that prior to the assessment of investigation they stopped the biology topic they were teaching and gave students practice through doing formative assessment (a mock examination) in a physics context that was very similar to the assessed task. The teachers then provided feedback to the students on how they could improve. This they justified by saying that they were ensuring their students were not disadvantaged because they were doing a biology topic whereas the assessment was set in a physics context.

Training for assessment involved an emphasis on what the students needed to write to achieve a particular grade, a practice noted also in the study by Cleaves & Toplis (2007). NCEA grades require a student to be able to describe their investigation to get an Achieved grade, explain their answer to get a Merit grade, and discuss their results to get an Excellence grade.

In Abrahams & Millar's (2008) view and according to research findings by Roberts (2009), both conceptual and procedural understandings are needed to carry out science investigation. Instead of developing these two kinds of understandings, students in this study were trained to perform in a way that matched the narrow focus encouraged by the NCEA assessment requirements for science investigation.

Changes in teaching practice after the introduction of NCEA

Teaching of science investigation for year 11 changed after the introduction of the assessment of practical investigation for NCEA Level 1. Of the region's year 11 science teachers who had taught before the introduction of the NCEA, 83% reported a change in their practice of teaching science investigation after the introduction of the NCEA.

Some regional teachers reported that since the introduction of the NCEA, they had changed the number of investigations they did in year 11 science. Teachers offered several reasons for the change in practice. Whether they did more, the same, or fewer investigations, the main reason offered for the change in practice was the need to meet the assessment requirements; concerns related to student learning were much less a factor. Another reason given was that complete investigation, a requirement of the assessment, was time consuming and took up to three lessons.

In the study school, a major change in practice following the introduction of the NCEA was the greater time devoted to teaching fair testing types of investigation in the first half of the school year as build-up to AS1.1. In the second half of the school year, after assessment, teachers said they reverted back to their pre-2002 practice of committing less time to fair testing and more time to other types of investigation and practical activities.

Study school teachers highlighted two related changes in practice. Firstly, they emphasised that students should learn the vocabulary required for AS1.1, for example 'independent' and 'dependent' variable and the 'reliability of data'. These are crucial ideas for understanding science investigation but the focus was more on learning the vocabulary rather than ensuring a deep understanding of the ideas. Wellington (2005) refers to the need for students to build the bridge between 'knowledge that' (observed phenomenon), 'knowledge what' (remembering facts) and 'knowing why' (understanding the reason for phenomenon occurring) (p. 107). In this instance, students learnt that they needed to repeat the trial several times but did not know why they should be doing so.

The second change to practice by teachers as a response to pressure to improve student performance, was training the students to become familiar with the template used for AS1.1. In the study school a simplified template was developed for use in year 9. One limitation of this template approach for learning investigation is that it was designed for fair testing types of investigation.

A noteworthy current assessment practice for over 75% of the region's teachers, and all of the school case study teachers, was to carry out formative assessment in the form of a mock examination or trial run, a practice also observed in the study class. Teachers said they saw advantages and disadvantages in using formative assessment of this kind. They reported that students valued the feedback that would help them to improve their performance. Teachers saw this as outweighing the disadvantages of workload, marking, administration, and management because it was helpful to students. Formative assessment in the form of a mock examination also helped teachers determine if the intended learning outcomes were met. Teachers said that they gave feedback to students about what they could do to get an Achieved, Merit, or Excellence grade. Formative assessment, as applied by these teachers, was different from that described by Bell (2005) and Bell & Cowie (1999); these researchers stressed the importance of assessment 'during learning' that 'relies on teachers developing in their pupils an orientation towards learning as distinct from performance' (Cowie 2005, p. 3).

Reliability and validity issues with assessment of investigation

Reliability focuses on the accuracy or consistency of results across assessors and over time (Harlen 2005). For assessment of investigation, high reliability 'would entail students getting the same results all the time irrespective of when the assessment is carried out and who marks it' (Harlen, p. 246). Training the students to achieve in assessment may enable students to rewrite the same answers and get the same result if the same assessment task is used under the same conditions. Clearly, this would not be an appropriate indicator of student understanding of science investigation because it does not test students' understanding and behaviour on a range of different investigation tasks; such an approach might well produce high estimates of reliability (because of the consistencies built into the process) but the approach compromises validity. Some teachers in the study school said that even though a student may get an Achieved grade for AS1.1, they could not say if that student was capable of achieving it in a similar assessment. Validity can be increased by combining five to ten assessments of investigation from different contexts and by using a template and tightening the criteria (Gott & Duggan 2002). The high level of student achievement in AS1.1 in the study school and nationally (both 83%) suggests that the assessment task was comparatively easy for students in year 11. Another explanation could be that it was poorly implemented because students were trained and given too much direction.

Although the criteria have been tightened, the implementation, it appears, does not reflect this change for the most commonly used tasks for AS1.1. Both the task and marking schedule are available on TKI and are easily accessible to students. According to the New Zealand Qualifications Authority statistics, the same tasks have been used nationally for over eight years. Potentially, students can find out what the task is and prepare for it and write the expected answers indicated in the marking schedule to get an Achieved, Merit, or Excellence grade. In the study school, where assessment took place in three lessons spread over two weeks, students had ample time to find out specific information required and use it in their report before marking and feedback had occurred.

The assessment of science investigation as required by NCEA and implemented in the study school has had negative side-effects, including: encouraging a surface approach to learning; providing a narrow focus on fair testing types of investigation; teachers giving students training to perform in such an investigation; and teachers' narrow use of formative assessment and feedback. These negative side-effects highlight issues of consequential validity. The assessment of investigation as prescribed and implemented may be doing harm and is therefore open to challenge in terms of consequential validity (Crooks 1993, cited in Hall 2007).

It is clear that much of the preceding commentary on the validity and reliability of the science fair test investigation draws as much on themes in the relevant literature as on the data gathered here. What can be fairly said is that when constructed carefully, administered appropriately, and interpreted properly, assessment of science investigation provides an in-depth window into how students apply their knowledge and skills to carry out an investigation (Harlen 2005). The NCEA requirement of assessment of a single fair testing type of investigation using a tightly structured task is likely to have increased reliability because of the consistencies built into the teaching and assessment; however, constraints are clearly placed on validity because of the narrowness of what is assessed.

Conclusion and Implications

Year 11 science teachers changed their practice of teaching science investigation in response to the requirement of NCEA internal assessment of science investigation. Teachers mostly taught fair testing type of science investigation that was linear and sequential as required for assessment rather than teaching a variety of science investigations as required by the curriculum. They prepared their students through repetition, teaching the language required for assessment, and providing one opportunity for a trial run to give student feedback on how to improve their grades. The use of a template is likely to have made the assessment more reliable. This change in teacher practice has consequences for student learning of science investigation. Continuation of such practice is likely to result in students seeing science investigation as learning what they have to do to get a particular grade rather than developing an understanding that science is predicated upon investigation.

Acknowledgement

The authors would like to thank Dr Janet Davies and Dr Stephanie Doyle for their guidance and support during the course of this research. We would also like to thank the teachers and students without whom this research would not be possible. The authors would like to acknowledge the support from Susan Kaiser in the preparation of this paper.

References

Abrahams, I. 2011. *Practical Work in Science: A minds–on approach.* London: Continuum.

Abrahams, I.; Millar, R. 2008. Does practical work really work? A study of the effectiveness of practical work as a teaching and learning method in school science. *International Journal of Science Education 30(14)*: 1945–1969.

Allen, M. 2008. 'Now this is what should have happened….': A clash of classroom epistemologies? *Eurasia Journal of Mathematics, Science & Technology Education 4(4)*: 319–326.

Bell, B. 2005. *Learning in Science: The Waikato research*. New York: Routledge Falmer.

Bell, B.; Cowie, B. 1999. Researching teachers doing formative assessment. *In*: Loughran, J. (Ed.) *Researching Teaching*. London: Falmer Press.

Cleaves, A.; Toplis, R. 2007. Assessment of practical and enquiry skills: Lessons to be learnt from pupils' views. *School Science Review 88(325)*: 91–96.

Cowie, B. 2005. How do pupils respond to assessment for learning? *Curriculum Journal 16(2)*: 137–151.

Crooks, T. 1993. *Principles to guide assessment practice.* Higher Education Development Centre, University of Otago, Dunedin.

Gott, R.; Duggan, S. 2002. Problems with the assessment of performance in practical science: Which way now? *Cambridge Journal of Education 32(2)*: 183–201.

Hall, C. 1997. The National Qualifications Framework Green Paper: What future for the framework? *Annual Review of Education 7*: 29–58.

Hall, C. 2007. *Planning the assessment for programmes and courses: A guide for tertiary level educators*. Wellington, School of Educational Psychology and Pedagogy, Victoria University of Wellington.

Harlen, W. 2005. Trusting teachers' judgement: Research evidence of the reliability and validity of teachers' assessment used for summative purposes. *Research Papers in Education 20(3)*: 245–270.

Hodson, D. 1990. A critical look at practical work in school science. *School Science Review 70*: 33–40.

Hofstein, A. 2004. The laboratory in chemistry education: Thirty years of experience with developments, implementation and research. *Chemistry Education: Research and Practice 5(3)*: 247–264.

Hofstein, A.; Kipnis, M.; Kind, P. 2008. Learning in and from science laboratories: Enhancing students' meta-cognition and argumentation skills. Pp. 59–94 *in*: Petroselli, C.L. (Ed.) *Science Education Issues and Developments*. London: Nova Science.

Hume, A.; Coll, R. 2008. Student experiences of carrying out a practical science investigation under direction. *International Journal of Science Education 30(9)*: 1201–1228.

Hume, A.; Coll, R. 2009. Assessment of learning, for learning, and as learning: New Zealand case studies. *Assessment in Education: Principles, Policy & Practice 16(3)*: 269–290.

Kanari, Z.; Millar, R. 2004. Reasoning from data: How students collect and interpret data in scientific investigations. *Journal of Research in Science Teaching 41(7)*: 748–769.

Lunetta, V.N. 1998. The school science laboratory: Historical perspectives and centres for contemporary teaching. Pp. 169–188 *in*: Fensham, P. (Ed.) *Developments and Dilemmas in Science Education*. London: Falmer Press.

Lunetta, V.N.; Hofstein, A.; Clough, M.P. 2007. Learning and teaching in the school science laboratory: An analysis of research, theory, and practice. Pp. 395–441 *in*: Abell, S.K.; Lederman, N.G. (Eds) *A Handbook of Research on Science Education*. New Jersey: Lawrence Erlbaum.

Merriam, S.B. 1998. *Qualitative Research and Case Study Applications in Education.* San Francisco: Jossey-Bass.

Millar, R. 2004. *The role of practical work in the teaching and learning of science.* Paper presented for the meeting of high school science laboratories: Role and Vision. Washington, DC: National Academy of Sciences.

Ministry of Education. 1993. *Science in the New Zealand Curriculum.* Wellington: Learning Media.

Moeed, A. 2010. *Science investigation in New Zealand secondary schools: Exploring the links between learning, motivation and internal assessment in year 11.* PhD thesis Victoria University of Wellington.

New Zealand Qualifications Authority. 2005. Achievement Standard 1.1. Retrieved 24 June 2009 from: http://www.nzqa.govt.nz/ncea/assessment/search.do?query=Science&view=achievements&level=01.

New Zealand Qualifications Authority. 2006. Statistics for schools. Retrieved 29 October 2009 from http://www.nzqa.govt.nz/qualifications/ssq/statistics/nqf–stats.do?ch=3210&year=2007&nqfLevel=0&st=0&cg=0&la=30&dm=1178&pc=250

Patrick, H.; Yoon, C. 2004. Early adolescents' motivation during science investigation. *Journal of Educational Research 97(6)*: 319–330.

Roberts, R. 2009. Can teaching about evidence encourage a creative approach in open-ended investigations? *School Science Review 90(332)*: 31–38.

Roberts, R.; Gott, R. 2003. Assessment of biology investigations. *Journal of Biological Education 37(3)*: 114–121.

Roberts, R.; Gott, R. 2006. Assessment of performance in practical science and pupil attributes. *Assessment in Education: Principles, Policy & Practice 13(1)*: 45–67.

Simon, S.; Jones, A.; Fairbrother, R.; Watson, J.; Black, P. 1992. Open work in science: A review of existing practice. OPENS project (1992), King's College University of London: Centre for Education Studies. Retrieved 12 July 2008 from http://www.comune.torino.it/sfep/praise/dwd/documents/references.pdf

Tytler, R. 2007. Re-imagining science education: Engaging students in science for Australia's future. *Australian Education Review*. Victoria, Australia: ACER Press.

Wellington, J. 2005. Practical work and the affective domain: What do we know, what should we ask, and what is worth exploring further? Pp. 99–109 *in*: Alsop, S. (Ed.) *Beyond Cartesian Dualism*. Netherlands: Springer.

Woolnough, B.E. 1991. Setting the scene. Pp. 3–9 *in*: Woolnough, B.E. (Ed.) *Practical Science*. Milton Keynes: Open University Press.