Hume, A. & Coll, R. (2010). Authentic student inquiry: the mismatch between the intended curriculum and the student-experienced curriculum. Research in Science & Technological Education, 28(1), 43-62.

AUTHENTIC STUDENT INQUIRY: THE MISMATH BETWEEN THE INTENDED CURRICULUM AND THE STUDENT-EXPERIENCED CURRICULUM

Abstract

As a means of achieving scientific literacy goals in society, the last two decades have witnessed international science curriculum redevelopment that increasingly advocates a 'new look' inquiry-based approach to learning. This fresh perspective on learning through inquiry promotes student engagement in investigations that reflects authentic scientific activity. Despite this renewed emphasis in national curriculum polices, recent research findings indicate that this curriculum goal is not being realised in classroom practice. This paper reports on the nature of the student-experienced curriculum where secondary school students are learning under a national curriculum that is intent on promoting students' knowledge and capabilities in authentic scientific inquiry. Using a multiple case study approach, the study gives insights into the student-experienced curriculum based on classroom teaching and learning experiences. The research findings show that rather than fostering creative and critical thinking, classroom experiences are leading to student learning that is largely instrumentalist and superficial. Closer examination revealed that layers of curriculum interpretation from several 'sites of influence' both outside and inside of the schools have a strong bearing on the curriculum enacted by teachers and experienced by the students, and run counter to the aims of the national curriculum policy. These interpretations are resulting in classroom teachers delivering structured teaching programme that feature substantially didactic pedagogies. Over-emphasis on fair testing limits students' exposure to the full range of methods that scientists use in practice, and standards-based assessment using planning templates, exemplar assessment schedules and restricted opportunities for full investigations in different contexts tends to reduce student learning about experimental design to an exercise in 'following the rules'. These classroom realities have implications for students

understanding of the nature of authentic scientific inquiry and support claims that school science is still far removed from real science.

Introduction

To achieve curriculum goals linked to scientific literacy classroom-based scientific inquiry has re-emerged as an emphasis in science curricula over the last decade or two (Crawford 2007; Toplis & Cleaves, 2006). This new-look inquiry approach, termed authentic scientific inquiry, encourages the 'doing of science' by students where they have the opportunity to experience the procedural and conceptual knowledge required to carry out investigation in a manner that mirrors the actual practice of scientific communities (Atkin & Black, 2003). The justification is that through this authentic inquiry "learners can investigate the natural world, propose ideas, and explain and justify assertions based upon evidence and, in the process, sense the spirit of science" (Hofstein & Lunetta, 2003, p. 30). Students may then become enculturated into science in a manner that ultimately helps them develop an understanding and appreciation of the nature of science (Collins, 2004; Driver, Leach, Millar & Scott, 1996; Duschl & Hamilton, 1998; Powell & Anderson, 2002; Weinburgh, 2003). Such investigations can also serve to motivate students' interest and desire to learn science (Hughes, 2004; Jenkins, 1996), engender attitudes and dispositions associated with those of autonomous and self-motivated learners (Deboer 2002; Reid & Yang, 2002), improve students' thinking and learning capabilities (Duggan & Gott, 2002; Haigh, 2003; White & Fredericksen, 1998), and facilitate cooperative learning (Hofstein & Lunetta, 2003).

What is Authentic Scientific Inquiry?

Scientific inquiry, as practiced by scientists, has been described as a complex social practice that involves participants interpreting, negotiating and justifying their inquiry approach in order to build believable and plausible explanations about how the physical world works (e.g., Wallace & Louden, 2002; Hofstein & Lunetta, 2003; Sandoval, 2005). The precise nature of scientific inquiry is difficult to define because scientists investigate the natural world in diverse ways (McComas, 1998; Watson, Goldworthy & Wood-Robinson, 1999), and scientists themselves have varying perspectives on how they work (Wong & Hudson, 2009). However, in a survey of the mental and physical skills accorded the title of 'scientific process skills', Harlen (1999) did find agreement in the literature that these skills were; "in one form or another, abilities related to identifying investigatable questions, designing investigations, obtaining evidence, interpreting evidence in terms of the question addressed in the inquiry and communicating the investigation process" (p. 129). Since this set of skills could be considered generic across a range of knowledge domains other writers contend that these skills only become scientific when they are applied in the context of science and informed and guided by scientific theory (Atkin & Black, 2003; Hodson, 1992).

In authentic science, scientists are routinely presented with open-ended problems which Reid and Yang (2002) define as problems where there are no data, known methods or established goals. In such situations all these components have to be developed by scientists in order to address the problem. Successful open-ended problem solving then depends on the knowledge and experience held by the people involved, and their ability to draw on appropriate and relevant information.

Hodson (1992) contends that scientists do this intuitively, using their own personal theoretical constructs and tacit knowledge of how to do science. He describes how this 'art and craft', or 'connoisseurship', gives scientists the "capacity to use theoretical and procedural knowledge in a purposeful way to achieve certain goals' (p. 133), and believes this only comes with experience of 'doing science' in holistic investigations in many different contexts. Thus to mirror authentic scientific inquiry students need to engage in open-ended problem-solving opportunities in a variety of contexts where they have to: draw on their existing science ideas to analyse the problem; plan a course of action; carry out the plan to obtain information that they can analyse; interpret and evaluate to reach a conclusion; and finally communicate their findings in some form (Duggan & Gott, 1995; Garnett & Garnett, 1995).

From Intended Curriculum to Operational Curriculum

While national curriculums may now give a clear lead to schools and teachers regarding authentic scientific inquiry for their students, moving from policy document to the operational curriculum (i.e., that actually experienced by students in the reality of the classroom) is not necessarily a rational and linear process (Atkin & Black, 2003; McGee & Penlington, 2001b; Ministry of Education [MoE], 2002). According to Carr et al. 2001 it is more realistic to view national curriculum policy as "the start of a cascade of interpreted curricula" (p. 18) that influence what students experience and learn in the classroom. These influences emanate from sources, or 'sites of influence' (Carr et al., 2001; English, 1997; Knapp, 2002; Spillane, 2004, Toplis & Cleaves, 2006), within the educational context that support and promote particular interpretations of 'worthwhile' learning. These sites of influence are contexts or arenas of action where participants have shared understandings of concepts and ideas due to the shared social contexts. Examples of such sites of influence could include national curriculum policy statements; curriculum support materials such as commercial publications; government educational support services, including provision of professional development for teachers; national qualification authorities and the qualifications; school and community aspirations for the education of students; and teachers' beliefs and values about teaching and learning (Hargreaves & Fullan, 1992). These sites exert varying degrees of influence on the operational curriculum, and research indicates some influences like national qualifications may play a more significant role in shaping the student-experienced curriculum than others (e.g., Atkins & Black, 2003; Carr et al., 2001; Harlen & James, 1997; Harlen & Crick, 2003; Reay & Wiliam, 1999). The process of implementing curriculum policy is one of interplay between what a curriculum statement says and the various interpretations and emphases afforded it by supporting materials, agencies, schools and teachers (Knapp, 2002; Spillane, 2004). Added to this mix are the cognitive, social, and language processes that are occurring within the classroom environment which also impact on this student-experienced curriculum, along with decisions that students themselves make consciously about learning (Nuthall, 1997).

Not surprisingly, recent research indicates that classroom implementation of curriculum goals around authentic scientific activity and inquiry-based learning is proving to be a slow process and far from straightforward (Atkins & Black, 2003; Rennie, Goodrum & Hackling, 2001; Wong & Hodson, 2009). The literature suggests that the manner in which investigative science is currently being taught in schools world wide, and the nature of the student learning still bears little resemblance to authentic scientific inquiry (Crawford. 2007; Haigh, 2005; Hipkins &

Booker, 2002; Author, 2006; Toplis & Cleaves, 2006). Frequently the slow pace of change and the nature of classroom inquiry learning are being linked to teachers' personal beliefs and views about teaching and learning science, the extent of their knowledge about authentic scientific inquiry and pressures exerted from influences outside the classroom (Crawford, 2007; Rahm, Miller, Hartley & Moore, 2003). To create classroom learning environments that are genuinely inquiry-based teachers face challenges that require them to use new instructional strategies and technologies, manage classrooms differently, place new emphases on process rather than content and use less traditional approaches to assessment (Schneider, Krajcik, & Blumenfeld, 2005). Facilitating student engagement in authentic scientific inquiry places further demands on teachers since they need to have a deep understanding of such activity - there is extensive evidence to suggest that few teachers possess such knowledge (Rahm et al., 2003; Wong & Hodson, 2009). However, the demands that assessment for qualifications places on teachers, particularly those of reliability (Toplis & Cleaves, 2006), appear to take precedence over curriculum goals in many schools and undermine teachers' abilities to provide opportunities for students to carry out authentic scientific investigations. The findings from the case studies in New Zealand verify this trend.

Background to the study

The Science in the New Zealand Curriculum (SiNZC) was introduced in the 1990s as part of sweeping educational and curricula reforms (Lange, 1988; MoE, 1993) in line with international trends of detailed and mandated national curricula accompanied by some form of national standards describing concepts for students to learn (Carr et al., 2001). The current SiNZC comprises eight progressive levels of broad learning outcomes known as 'achievement objectives', which are grouped under achievement aims in six learning strands (Haigh, 1995). Bell, Jones & Carr, 1995 summarise the key content for students to learn as: investigation and problem solving; understanding scientific knowledge, understanding the nature of science and; understanding the influence of science on people. While the terms 'scientific literacy' and 'authentic scientific inquiry' are not specifically mentioned in the document, elements now accepted as related concepts do appear. For example, one of the six strands, the *Developing Scientific Skills and Attitudes* strand, deals specifically with scientific skills and attitudes, but these are also integrated into the content of achievement aims and objectives of the other strands. The curriculum policy statement requires that both content and process skills are to be given careful consideration and attention in teaching, with practical work, in particular investigations, cited as a vehicle for developing scientific understanding.

> Investigations provide key opportunities for students to extend their understanding in science. They also enable students to develop the scientific skills and attitudes required to enhance their ability to explore phenomena and events and to solve problems. It can be expected that, as they learn, students will show an increasing sophistication in the skills they use in their investigations (MoE, 1993, p. 42).

At Levels 5/6 (i.e. Years 9-11) of the *Developing Scientific Skills and Attitudes* strand this theme of inquiry has been translated into achievement objectives that focus on a particular type of investigation, by specifying that students "design fair tests, simple experiments, trials and surveys, with clear specification and control of likely variables" (MoE, 1993, p. 44). It is at Level 6 of the SiNZC (Year 11) that achievement objectives are first assessed for national qualifications – the National Certificate of Educational Achievement (NCEA) which was introduced in 2002.

The essential building block of NCEA is the *achievement standard*. These standards are statements, in the form of generic performance criteria, which describe what students need to know and do in order to gain credit. Science achievement standards are designed to assess learning as defined by the achievement objectives in the SiNZC, and judge student performance at four levels (non achievement, achievement, achievement with merit and achievement with excellence). At least half the achievement standards for conventional school subjects are externally assessed, to address concerns about issues to do with internal assessment, such as moderation and teacher workload (Lee & Lee, 2001).

The standard of particular relevance to this inquiry is Science Achievement Standard 1.1 *Carrying out a practical investigation with direction* (SAS 1.1). This standard, which has since been revised (see later), stated that student investigations should be based on situations arising from content drawn from achievement objectives up to Level 6 (Year 11 students) of the SiNZC (MoE, 1994). An investigation was defined in the standard as an activity covering the complete process from planning to reporting, and was to involve the student in gathering primary data (i.e., generating and recording their own data). Under direction from the teacher, students were expected to: produce a workable plan, containing a purpose, provision and evidence of trialling, key variables and how they will be controlled, a method for data collection and consideration of factors such as sampling, bias, sources of error and sufficiency of data;

- execute the plan, collect appropriate data and record in a table or other systematic way, and process to establish a relevant pattern or trend. Data processing is expected to usually involve calculations such as averaging;
- interpret the processed data in relation to the purpose of the investigation and;

- write a report following written guidelines from the assessor. Sections of this report are to usually include the purpose and final method used, recorded and processed data showing links, interpretations, a conclusion linking findings to the purpose, and an evaluation or discussion.

The standard also specifies the use of a format for student reporting of the investigation, and the MoE provides templates for planning and reporting via exemplars (e.g., MoE, 2003). Various publishers quickly produced texts based on this achievement standard when NCEA was first implemented. They ranged from textbooks providing content and exemplars (e.g., Hannay et al., 2004), to student laboratory manuals (e.g., Author, 2002). Most tended to closely follow the format and requirements of the standard and accompanying exemplars in their interpretation of carrying out scientific investigations.

In a study that investigated teachers' perceptions of changes they have made in the delivery of their Year 11 science programmes since the introduction of NCEA in 2002 Hipkins (2004) found the teachers were already adapting their classroom practice to meet NCEA requirements. In the assessed inquiry-based component of their courses for SAS 1.1, some of the teachers had quickly formed the view that student investigations "had to be focused within the narrow, formally presentedframework of the reporting schedule for this achievement standard" (Hipkins, 2004, p. 9), and were redesigning their programmes accordingly.

Significance of the Study

Reviews of the international literature reveal that most current understanding about the nature of student inquiry learning comes from large-scale reviews of research and meta-analyses of international literature (Carr et al., 2001; Hipkins et al., 2001), often based on evidence obtained through surveys of teacher and student perceptions of classroom practice. On the other hand, detailed classroom-based case studies are relatively rare and considered important since they facilitate direct study of the interplay of the more intricate and specific variables of each classroom environment, such as teacher expertise and student interactions (Jones & Baker, 2005). The case studies reported in this current paper reveal the educational reality for students who are experiencing a classroom curriculum, guided by national curriculum goals of authentic scientific inquiry and assessed by a national, standards-based qualification – an educational situation that is becoming common internationally. Research findings that can shed light on the match between curriculum intent and classroom reality in such educational environments are important when evaluating how effectively curriculum goals are actually being met, and help inform decisions about what steps may be needed to improve outcomes for students. The case studies took place in the context of a national science curriculum (Ministry of Education [MoE], 1993) that sought to promote students' engagement in authentic inquiry. Each study involved a Year 11 science class where students (15-16 year olds) were learning how to perform investigations for Science Achievement Standard 1.1 *Carrying out a practical investigation with direction* (SAS 1.1) towards their National Certificate of Educational Achievement (NCEA). The Year 11 course was chosen because for many of the students this class was their last opportunity for formal schooling in science, and likely to be a time when they formed lasting impressions of the nature of scientific inquiry. The ideas and beliefs they form during their classroom investigations could have implications for their scientific literacy as future citizens, especially the extent to which they understand and appreciate the ways scientists work to produce scientific evidence, solve problems and build knowledge.

The case studies were set in two large New Zealand secondary schools, *River Valley Boys' High School* and *Mountain View High School* (pseudonyms). Both school populations were similar in that they were predominantly of New Zealand European ethnicity (77% and 68% respectively) and each had 12% Māori (i..e, people who self-identify as indigenous New Zealanders). *Mountain View* also had a significant proportion of Asian students (15%). Each case study involved a female teacher, and four to five Year 11 students (15-16 year olds) who were studying SAS 1.1 towards their NCEA qualification. The students at each school were in classes representing a very broad band of mid–range of academic abilities based on achievement information from common internal assessments, such as school exams and tests, from the last two years of their compulsory schooling in science i.e., Years 9 and 10. This broad band – approximately 80% of the whole Year 11 cohort. The remaining 20% comprised students at either ends of the ability spectrum i.e., students with learning difficulties at one end and those with special abilities at the other. At *River Valley* Jenny (pseudonym) the teacher held a master of science degree in genetics and was in her eighth year of teaching. In her interviews she declared that practical work played an important part in her pedagogy, because she believed strongly that 'hands-on' activity helped learning:

> I'm very practical orientated, I suppose as a science teacher. Most of my classes will have a practical everyday if at all possible. I think that's what helps them learn in science hands-on … certainly not by me standing up in front of the classroom and talking.

When asked if she viewed practical work as part of investigative science she responded affirmatively, citing instances where she gave students the opportunity to carry out practicals without initial teacher direction:

> Often with classes I won't even introduce something. They'll just go ahead and do the practical, find out what's happening. I'll go around and interact with them quite a bit and sort of give them leading questions the whole way through, then we'll come back and discuss it.

She considered these practicals were investigations, in the sense that she took a more "hands-off" approach. The all-male student group included Martyn, Peter, Mitchell, Eddie and Sam (pseudonyms) who were all New Zealand Europeans.

In contrast at *Mountain View* the teacher Kathy (a pseudonym) had begun her teaching career three years earlier at *Mountain View* after completing a conjoint bachelor's degree in science and teaching. In articulating her views of teaching in her first interview Kathy did not volunteer many opinions about the nature of science education or science, but she did view science as context that offered opportunities for the use of teaching strategies that resulted in effective student learning. For example, in her teaching approach she valued the hands-on learning that science experiences could offer:

> I try and use lots of different techniques and activities with them to help students. We are lucky in science where we can do heaps of hands on … I am doing practical work all the time for that reason because most of them seem to learn better that way.

When asked later in the interview what she believed good investigatory science to be, she considered this for some time before replying "Well I guess it is all about coming up with a question and trying to prove or disprove it*"*.

The mixed-gender student group Anne, Carol, Alex, Mark and Steve (pseudonyms), came from different ethnic backgrounds. Anne, Mark and Steve were New Zealand Europeans, Alex a New Zealand born Asian student with English as his first language, and Carol a recently emigrated Asian student for whom English was a second language.

The three research questions for this inquiry were:

- *What* science are New Zealand science students learning in NCEA classroom programmes for SAS 1.1
- *Why* and *how* are New Zealand science students learning the science they learn in NCEA classroom programmes for SAS 1.1?
- What match is there between the intended curricula (i.e., those of the SiNZC and the teacher) and the operational science curricula (i.e., those experienced by New Zealand science students)?

Theoretical Underpinnings

This inquiry was conducted within an interpretivist paradigm, drawing sociocultural and linguistic perspectives into a cognitive constructivist model of the development of thinking processes as a framework for enhancing understanding of learning (Nuthall, 1997). These teaching and learning theories were chosen to underpin the study because they have in common the view that knowledge has to be personally experienced. In a constructivist-based view of learning, students experience changes in what Leach and Scott (2003, p. 92) term the "mental structures" of individuals, that is, their concepts, schema or mental models. Individual learners construct their own knowledge motivated by the need to make sense of experience in light of their existing understandings. The sociocultural stance on learning is that "thinking and learning are not seen as an activity of the mind in isolation, but rather as part of, or constituted by, the visible social interaction that takes place between members of a community" (Nuthall, 1997, p. 701). What counts as knowledge is situated in the practice of that particular community and defined in social interactions (Barnett & Hodson, 2001; Black, 2001). The linguistic perspective acknowledges the acquisition of language as a semiotic process (i.e., one of making meaning) that is central to all learning. Language is the means by which concepts are introduced and discussed by learners on the social plane, and the tool for individual thinking once concepts are internalized (Leach & Scott, 2003).

Methodology and Methods

The interpretivist-based methodology (Guba & Lincoln, 1989) comprised a multiple case study approach utilising qualitative research methods of unobtrusive observation, semi-structured interviews and document analysis. In the first of two case studies a total of 12 one-hour lessons were observed and six interviews conducted (3 with the teacher and 3 with the students as a group), while in the second case study fewer lesson were observed (7 in total) and 5 interviews with participants were conducted (3 with the teacher and 2 with the students as a group).

Classroom observations in both case studies were recorded via field notes and audio-taping, and the interviews were audio-taped and transcribed verbatim. The interview transcripts were sent to interviewees for verification and alteration (if desired by the interviewee). Audio-tapes of the classroom lessons proved difficult to transcribe accurately because of background noise and the students tending to mutter at times, but they did add to the general pool of data collected in each case study, often corroborating other data. A variety of documentary material related to the teaching and learning occurring in each case study was examined, including departmental guidelines for teachers, textbooks and student workbooks, notes and assessment items such as exemplars and student scripts. The case study approach was used in order to facilitate a holistic, interpretive investigation of events in context with the potential to provide a more complete picture of the science curriculum students were experiencing compared to other modes of research (Adelman, Kemmis & Jenkins, 1980; Bell, 1999). Students were viewed as intentional participants in classroom activities and the interpretive analysis concentrated on their perspectives of classroom reality.

The constructivist, sociocultural and linguistic teaching and learning theories underpinning the study were initially surveyed to find suitable parameters on which to base data collection decisions. Working definitions for the *what*, *why* and *how* of student learning were devised, and guided the analysis, interpretation and discussion of the findings. The *what* of student learning, for example, was defined as those scientific concepts, skills and procedural knowledge students were acquiring and demonstrating through their words and actions during teaching and learning episodes ((Bell, 2005; Duggan & Gott, 2002; Hodson, 1996; Skamp, 2004)). Instances of *why* students were learning were characterised by the circumstances that led to students achieving that learning; what learners did in order to learn was the key parameter chosen to define the *how* set of data. This approach to the *how* of learning focused on the thought processes and actions attributed to students as they learn.

To enhance the trustworthiness of the inquiry process (Guba & Lincoln, 1989) particular attention was paid to strategies that would maximize the quality of data gathering and processing within the constraints of the study. For instance, a decision to take a two-case study approach in this study helped to mitigate the impact of some limiting sampling factors (i.e., small sample size in each case study) and promote transferability by allowing findings from two case studies to be compared and contrasted. Triangulation, of both data collection and sources of data, sought to promote the dependability, confirmability and credibility of the study by reducing the likelihood of researcher bias (Erickson, 1998) and producing sufficient wealth of evidence to allow a high degree of convergence (Bell, 1999; Keeves, 1998). Prolonged and extensive observation helped to establish the dependability of the data (Spindler & Spindler, 1992), along with detailed auditing of the inquiry process and respondent validation of the raw data, which also endorsed the confirmability of the data. Rich descriptions of the case studies, using narrative style, were used in places to allow the participants voices to be heard (Bishop, 1997), and to help to build an inferential bridge to those other groups to whom the findings may be applicable (Shulman, 1981). The analytic categories of *what*, *why* and *how* also enhanced the comparability and transferability of the research findings by giving readers greater opportunities to make meaningful comparisons with their own situations (Goetz & LeCompte, 1984).

Observations from Classroom Sessions: An Overview

At both schools many decisions to do with classroom practice were not made by the individual teachers, but were made collectively at departmental level in the form of departmental guidelines. These guidelines were based on recommendations, including exemplary materials, from the New Zealand Qualifications Authority (NZQA) – the agency who administer the NCEA qualification - which departments and classroom teachers were obligated to follow under school accreditation requirements. Thus at both schools the content of departmental guidelines was very similar, and both case study teachers adhered closely to departmental guidelines in their teaching and learning programmes.

At *River Valley* the teaching and learning took place during 12 one-hour lessons over a three-week period, late in term one of a four-term year. In contrast, students at *Mountain High* experienced a staggered teaching and learning programme, 11 hours in total. Their teaching started with five one-hour lessons late in the first term, followed by a three-week break before another four consecutive lessons early in the second term. Two weeks later *Mountain View* students attended a single timetabled session (two hours) within the school's mid year internal exam programme where they underwent the formal assessment for SAS 1.1. Despite the variation in the overall timing and duration of the teaching and learning sessions at the two schools, the sequence of lessons in both schools showed strong parallels. Each sequence could be divided into three distinct phases: the preparatory phase (instructional sessions); the practice phase (the formative assessment); and the formal assessment phase (the summative assessment).

The Preparatory Phase

In this first phase students in both classrooms were introduced to the requirements of SAS 1.1 and key concepts and skills associated with investigating relationships between two variables. .

> In the initial instructional session Jenny informed the students that the standard was going be taught in a block, including a formative assessment where they would be "learning how to do it", followed by a "revisit, then a final summative at the end of Term 1". She then introduced students to the standard through a whole-class, guided reading session of the first two pages of the student workbook (i.e., Author, 2002). These pages provided the achievement criteria from the standard and explanatory notes elaborating the meaning of key terms in the criteria like 'purpose', 'workable plan' and 'sources of error'. Students took turns to read sections aloud and Jenny punctuated the reading with questions (mostly closed questions to do with recall and procedure) and further clarification, placing emphasis on the meaning of terms, aspects of experimental design, and what was needed to achieve particular grades.
>
> (field notes and audio-tape transcripts, *River City*, Lesson 1.
> Note text inside quotes represent direct quotations for audiotpes)

Lesson content in these largely instructional sessions focused on: terms, definitions and procedures to do with fair testing: specific skills such as making observations and measuring, tabulation and averaging of data, plotting graphs and the planning and reporting of fair tests using templates and; how to meet the assessment requirements of SAS 1.1 as depicted in assessment schedule exemplars. Less time was devoted to the first phase at *River Valley* (three lessons compared to five at *Mountain High*), and Jenny also revised specific science concepts that featured in the investigation her students were to perform in phase two (rates of chemical reaction and preparation of solutions of given concentration by dilution).

The Practice Phase

In the second phase students at both schools participated in a mock assessment known as the 'formative assessment', designed to give students practice at performing a whole investigation under test-like conditions. In the following excerpt Kathy at Mountain High had first asked students to individually design a plan to investigate the rate at which their magnesium metal reacts with hydrochloric acid. She had then taken in the individual plans, read them and on the basis of this assessment placed the students into mixed ability groups. Thus the research group students were dispersed through the various other groups.

> She instructed each group to talk amongst themselves, share plans and then write a shared or common method that they all agreed on. If there is no change in a particular plan the students could write on their script "see original plan". Kathy gave some last minute advice: "Be careful you have only 10cm of magnesium and a set volume of hydrochloric acid. Think carefully about your quantities. It needs to be written in your method". She ran through the report template, reiterating key points: "Every person has to write their results in … write the group method. Everyone has to have a copy of it … do the experiment as a group … record data in a table". The students had a third class period in which to finish the report. They could attempt to write the report up in this second practical session, but Kathy insisted: "Don't rush it".
>
> (field notes and audio-tape transcripts, *Mountain High,* Lesson 6)

Again there were many commonalities between the two case studies:

- the mock assessment took place over four lessons, with each lesson covering in turn, the planning, data collecting, reporting and feedback stages of the investigation;
- the science context for the investigations was the same (both teachers used the same exemplar materials for investigating the effect of factors such as temperature or concentration on the rate of reaction between magnesium metal and hydrochloric acid); students worked in teams of four for planning and data gathering but as individuals for the reporting;
- the format, timing and reporting requirements of the mock assessment activity closely matched those of the summative assessment in phase three; and
- teacher direction was highly evident, including extensive and targeted feedback for students related to the assessment schedules for the task.

In addition, at *River Valley* students initially peer assessed each other's reports using a common assessment schedule and provided verbal feedback to one another before the teacher provided global feedback to the class.

The Formal Assessment Phase

In the third phase for their formal assessment, known as the 'summative assessment', students again performed fair test investigations in groups along similar lines to the practice investigation in the second phase. They initially planned as individuals, then collaborated as a group to produce a single plan and obtain data, and finally wrote up the reports individually. The planning and reporting templates were virtually identical in the two schools, however, the science contexts for the investigations were different. Students at *Mountain Valley* performed their investigation in the context of reaction rates again, this time the relationship between surface area and the rate of reaction, while students at *River Valley* performed their investigation in the context of pendulums which they had no prior experience of in the course. Students at *Mountain View* planned and executed their investigation with relative ease, whereas the study group at *River Valley* experienced difficulties carrying out their plan investigating the relationship between the length of a pendulum and its period, that is, the time taken to complete a full swing. They were unable to operate the pendulum successfully and consequently could not record sufficient data. However, they were very savvy of assessment techniques and showed adeptness at 'playing the system' as the following excerpt shows:

> Within the closing stage of the practical session the group scrambled to complete and record sufficient runs for their data processing and interpreting phase. The four group members frequently interchanged roles as they each took it in turn to record their own copy of the results (which they needed for the write-up in the following session.). All other groups had finished their data collection and were listening as Jenny covered points for the write-up. Martyn, Peter, Mitchell and Eddie continued operating their pendulum and consequently missed hearing what Jenny was saying during her briefing. In their rush to finish confusion set in: "Is this the third or fourth one?" asked Mathew who was recording and calculating. When the pendulum continued to collide with the support arm Peter commented, "You'll have to estimate", while Eddie was convinced they should "make up the rest". Mitchell agreed: "Lets make up the rest, and take 16 seconds as the average" and Martyn confirmed, "It will still give us our results". Each group member had a complete set

> of written data by the end of the practical. Jenny allowed the class to view the background science notes before the end of the period before collecting in all papers to retain overnight.
>
> (field notes and audio-tape transcripts, *River City*, Lesson 10)

At the last minute the students resorted to recording their remaining results from non-existent data and then used these fabricated results to complete the reporting section of the assessment.

One other significant difference between the case studies in this formal assessment phase is that, unlike the students at *River Valley,* the five students in the *Mountain View* study did not work in the same groups for the summative investigation. Kathy purposefully decided groupings for the summative assessment at Mountain High on the basis of results from the formative assessment, so that each group intentionally had at least one student who had demonstrated advanced investigative capabilities.

What were Students Learning About Scientific Inquiry?

Findings from both case studies indicated that the learning many students were achieving about scientific inquiry closely matched that which their teachers intended them to learn. The content of the teachers' intended curricula is summarised in Table 1 below and represent a synthesis drawn from data collected during teacher interviews, observation of classroom lessons, departmental guidelines and notes, and student workbook and text (Author, 2002; Hannay et al., 2002). The summary in Table 1 serves to indicate the key concepts and skills that the majority of students did demonstrate familiarity with during their investigations, and the following sections elaborate further on the extent and nature of this learning.

(Insert Table 1 here)

Data gathered from classroom observations of student and teacher actions, interviews and assessment information found that the science students were acquiring in the teaching and learning programmes of both case studies was very similar in most respects, and linked to one particular form of scientific investigation – fair testing. Students' classroom experiences focused on investigating cause-and-effect relationships between physical phenomena, and their thinking and learning revolved around how to plan, carry out and report the findings of fair tests into these relationships. These findings indicated that in both classroom settings the students progressively learned concepts and skills about science investigations that reflected those broadly defined in the NCEA Science Achievement Standard 1.1 *Carrying out a practical investigation with direction*, and those specifically required to complete the generic planning and reporting template for assessment tasks provided by NZQA and the Ministry of Education (MoE). This knowledge and skills were reinforced by published texts (Cooper et al., 2002; Hannay et al., 2002) that were used by students.

This interpretation of the findings is based on the detailed match between the nature of the student learning evident in the findings and the content of various assessment tasks, particularly the assessment schedules, provided by the MoE and commercial text for exemplary and practice purposes in classrooms. Students' oral and written language, for example, showed increasing use of terminology associated with fair testing and understanding of the protocols that were prominent in these NCEA materials, such as: the independent variable and changing it systematically; the dependent variable and repetition of measurement; processing data and line graphs and lines of 'best fit'; and interpretations of data as findings related to the purpose of the investigation. The findings suggest it was the assessment schedules of these tasks that effectively prescribed the concepts, vocabulary, skills and procedural knowledge students were gaining during the investigation exercises.

Strong indications of the influence of assessment tasks and their schedules on students' learning also came from examination of the fair test plans that they produced as part of their investigations, and observation of their actions in implementing these plans. In exam-style sessions in the summative assessments, all students were able to individually produce plans that varied in quality from feasible (could be workable but lacks a few details) to workable (could be followed independently without further clarification). The minority of students who produced workable plans had identified and controlled key variables, and described means of obtaining, recording and processing relevant, accurate and reliable data if the plan was carried out as written.

It is important to note that there was little evidence of open-ended planning and investigation when students were required to do full investigations. All plans closely adhered to the experimental design inherent in the planning template used for assessment tasks and teachers had given considerable direction and support to the students about the content of these plans prior to planning sessions. This teacher direction provided the particular scientific relationships to be investigated and relevant experimental skills and techniques, even to the point of identifying the independent and dependent variables to be investigated in one of the case studies. The students thus went into planning sessions, even for the summative task, well informed about the procedural knowledge needed for that investigation. However, there were differences between the two case studies in the depth of understanding and level of experience with the background science concepts that students brought to the summative planning sessions, which impacted on their abilities to link their findings with science concepts.

As in the formative investigations, the students in both case studies worked in groups for the practical summative sessions with the result that all students had the opportunity to access workable plans. Both case study groups produced data from their experimental work, which allowed them to continue the processing, interpreting and reporting aspects of the investigation. However, since the summative practical session was unable to be observed by the researcher in one case study the students' performance from the two case studies could not be compared in this aspect of the investigations. In the case study where the researcher was present during the practical work the ability of students to collaborate successfully in the refinement and performance of their plan in the summative investigation was observed to falter at times. These students had gained an appreciation of the need for trialling to gauge the workability of their plan and experimental methods from their formative experiences, but the group focus on decisions to do with technical details occurred at the expense of decisions to do with method. Hasty last minute decisions about procedure ultimately proved to be costly. In addition, some critical logistical points were overlooked like task delegation, and as a consequence planned decisions were not always adhered to in the summative assessment. These student actions suggest that their understanding of some of the finer points of experimental design, such as repetition and appreciation of the depth of forward planning needed were superficial, and their level of experience and expertise with the technical components of this experimental context (the pendulum) limited. Despite these ‘procedural hiccups’ the students knew what data would be sufficient to allow them to accomplish the rest of the task, and they made a pragmatic decision to ‘cook their results’. In this sense they did not achieve their teacher’s intended curriculum of ‘good science’ by fabricating results, but they did demonstrate an understanding of how to effectively meet assessment specifications. This action gave each group member access to a set of seemingly valid and reliable data, which they subsequently recorded and processed in their individual written reports. As can be seen from the excerpt below students achieved success but the investigation was far from authentic.

> Peter was awarded achievement with merit for his summative grade, and he showed evidence of learning in all three sections of his report. His very detailed step-by-step method, which now included a list of equipment and the averaging of results, earned him excellence for first section. In his results section Peter failed to meet several criteria necessary for merit, including mention of the independent and dependent variables in his graph title and the line of best fit - he had attempted a curve, which was not the best fit for his data. However, because he had met so many of the excellence criteria, such as repetition (6 repeats for each pendulum length), and the use of average results Jenny took a holistic approach to marking this part and awarded an ‘on balance’ grade of merit. Teachers in the department had been given the discretion to use their judgement in this manner for situations like this. Peter’s results table now

> showed units outside the table, which was an improvement on his formative exercise. His interpretations and conclusion statements demonstrated a much deeper appreciation of what was required in this section of the report. He made appropriate interpretations based on his data and had partial success explaining his results in terms of the background science, including the roles of gravity and kinetic energy, and the relationship between the length of the pendulum, distance travelled and the period length: "Science ideas that explain the trend in the results link back to the force of gravity, from the horizontal position, gravity pulls the bob down. The shorter the string, the faster time it will take because it has less travelling distance". However, his next phrase shows an. alternative conception when he goes on to explain that the bob on the shorter string "picks up speed faster which is why the shorter the string the quicker the period time". Peter also evaluated their experimental work, and identified appropriate amendments to the method, like changing to a "spherical bob", but like Martyn he did not divulge the group's fabrication of results. He was awarded excellence in this section.
>
> (Student Assessment documents, *River City)*

In their written reports all students in the case studies individually recorded and processed their data sufficiently accurately enough to identify a relationship between the independent and dependent variables. Most students in Case Study A made correct data processing decisions in choosing to use line graphs, but several had trouble correctly drawing a line of best fit. Key table and graphing features that were missing from most their formative scripts were addressed in their summative scripts. In contrast, in Case Study B some of the finer details of data recording protocols for tables were missing, and all students graphed their data using bar graphs, instead of the more appropriate line graphs for identifying cause-effect relationships between two variables. Students in both case studies were able by the close of the summative assessment to draw conclusions based on their findings and related to the purpose of the investigation. Generally speaking, however, few of the students demonstrated the capacity to fully interpret and explain their results by linking their findings to existing science concepts, and while some students were attempting to evaluate the robustness of their findings even a very able student Alex at Mountain High only managed a superficial critique of his methodology. Alex had linked his findings to the relevant science background but:

> his evaluative comment was restricted to one sentence: “If I were to do my experiment again I would make sure that I dried the beakers out after I washed them each time, which would make my experiment more accurate”. He did not explain why this action would make the experiment more accurate.
>
> (Student assessment documents, *Mountain High*)

In the *River City* case study lack of familiarity with the background science in the summative task hampered students in their ability to link their findings with theory, to the extent that even Martyn an able student who had some success with this aspect in a formative task could not succeed in the science context of the summative task.

> In the last section he (Martyn) produced acceptable interpretations and conclusions about his data and made a good attempt to improve upon his evaluation. He recognised the group’s difficulties with the equipment and

> finishing within the allocated time, but did not identify or acknowledge that these difficulties may have been caused in part by the lack of effective teamwork: "The problem with this experiment was time. We ran short of time, which made us more hurried and less careful. Because of this the experiment was slightly less accurate". Martyn did not mention the 'approximations' they had made with their data due to running out of time. In attempting to link the observed trends in the data to the science ideas behind the pendulum, he was unable to use the science concepts to explain why the pendulum length affects the pendulum period. He instead tried unsuccessfully to explain why the pendulum swings in an arc. For this final part of the report he received a merit grade, and achievement with merit for the overall standard.
>
> (Student assessment documents, *River Valley*)

Why **and** ***How*** **were Students Learning?**

Interviewing the students and their teachers, observing them interact in class, and examining support materials and student records revealed w*hy* and *how* students learned about fair testing and the assessment requirements of the AS 1.1 were direct consequences of three influences: the content of their teachers' intended curricula, the pedagogical approaches and techniques that their teachers used, and the learning strategies that students employed. The key findings sourced from data obtained in classroom observations, interviews and documentation are summarised in Table 2 below.

(Insert Table 2 here)

At first glance, since the findings in both case studies show a close match between the teachers' intended curricula and that experienced by students, the most obvious reason *why* students in these case studies learned about fair testing in science and assessment procedures for NCEA, is that the teachers made the decision to deliver this particular content in the teaching and learning programmes. These decisions meant the teachers' instructional intentions focused on concepts, skills and procedural knowledge to do with investigating cause-effect relationships between variables and meeting the assessment requirements of the achievement standard. Clearly if certain content was not included by teachers in their programmes, then the likelihood of this 'extra' knowledge being accessed by students via classroom teaching was limited.

Close examination of the findings shows that these teacher decisions about lesson content were influenced most directly by their respective school departmental guidelines for delivering Science Achievement Standard 1.1 *Carrying out a practical investigation with direction*. These guidelines were, in turn, based on materials (planning templates, and exemplar assessment tasks and schedules) provided by the MoE and NZQA to support teaching and learning programmes for the NCEA Science Achievement Standard 1.1 *Carrying out a practical investigation with direction.* Teachers' classroom curriculum planning decisions were not directly influenced by the specific requirements of the SiNZC as stated in the document, but more by interpretations of that national policy by NZQA and school science departments. Similar interpretations to NZQA of the SiNZC requirements were also promoted by other sites of influence, including teacher professional development providers and support agencies, and publishers of textbooks. The many similarities between the students' experienced curricula in the case studies, the teachers' intended curricula, and the science content promoted by assessment support materials and providers with the learning measured by Science Achievement Standard 1.1 *Carrying out a practical investigation with direction* provides a strong indication of *why* the learning these Year 11 students achieved in the case studies focused on fair test investigations and assessment procedures. It emerged that the science assessed by the Science Achievement Standard 1.1 *Carrying out a practical investigation with direction* was chosen by teachers, curriculum support agencies and textbook publishers as the basis for the content of the curricula they delivered.

Discussion

This study sought to gain some insights about the possible nature of the student-experienced curriculum as our Year 11 students learn about scientific inquiry from the perspectives of some actual teachers and students in the classroom. The evidence emerging from this present interpretive study into a student-experienced curriculum demonstrates strong parallels between *what* students were learning about scientific investigations in these New Zealand classrooms and those learning trends identified from the international literature. In the case studies *what* students came to perceive and experience as scientific investigation was the single, linear and unproblematic methodology of fair testing. Nevertheless, within a narrow context of fair testing, the students did manifest many of those physical and mental skills generally agreed upon in the literature as the ‘scientific process skills’ (Harlen, 1999). For example, they were able to produce appropriate scientific investigations, obtain relevant information, interpret evidence in terms of the question addressed in the inquiry and communicate the investigation process. However, with the support of planning templates, exemplar assessment materials and teacher direction the standard simply required students to follow a set of rules and procedures which they learned in practice assessments rather than coming up with original solutions to experimental design. The planning templates students used in these case studies served as blueprints, in effect restricting *what* students were learning about planning to following a formula, just as Roberts and Gott (2002) observed happening for students performing investigations under similar assessment conditions for Science Attainment Target 1 of the National Curriculum in Britain. Consequently in these case studies, students' ability to identify investigatable questions was not evident, simply because the nature of the teacher direction and the structure of the planning template did not require students to identify ways in which their scientific understanding could be expanded via investigation. The purpose of the investigation was a 'given', and students' only task was to craft a question specifying the cause and effect relationship they were investigating – in fact, in one of the case studies the teacher' direction even extended to identifying the independent and dependent variables for students in her lead in comments to the assessment. Students were not participating in authentic open-ended investigations, where they had the responsibility for determining the purpose of the investigation and the question to be investigated, as 'real' scientists would (Hodson, 1992; Reid & Yang, 2002).

Some authors comment that in terms of *what* students learn about carrying out authentic scientific activity, students tend not to learn to take account of scientific theory in planning their investigations and interpreting their results (e.g., Atkin & Black, 2003; English & Wood, 1997; Hodson, 1992). That claim is not fully supported by the findings of this study because students for the most part did take account of some scientific theory in the performance of their investigations. For example, they called on their prior learning of scientific concepts, skills and procedural knowledge related to fair testing and the scientific context to complete their planning template and conduct their investigations. The more able students also made some valid links between their findings and their scientific existing understanding, but in problem-solving situations that Reid and Yang (2002) would define as more closed in nature than open since teachers gave the students substantial guidance with the goals of the investigation, the scientific background and the procedures to be used. As a result of these student investigations there was little evidence that these activities were generating conceptual change for them as learners (English & Wood, 1997). The practical work appeared to be serving more illustrative purposes by reinforcing rather than expanding students' existing scientific understanding. In this sense, while students were practising skills and gaining experience with procedural aspects of fair testing students were not engaging in activity that reflected authentic science investigation.

So for most students in these case studies *what* they learned about scientific investigation was confined to applying a 'set of rules' about fair testing, to illustrate and confirm scientific concepts covered in the instructional part of their classroom programme and to meet assessment requirements. The use of templates and exemplars in the teaching and learning programme produced the 'seen exam' phenomenon described by Roberts and Gott (2004), providing the required protocols for assessment success, and not requiring students to demonstrate the sort of tacit, intuitive knowledge in their science investigative abilities that comes with wide experience and understanding (Hodson 1992), such as creative thinking in experimental design. Students' learning was characterised by lower to middle order thinking (Bloom et al., 1956; Anderson & Krathwohl, 2001), with only a few able to display some higher order critical thinking skills. The nature of the learning for most students tended to be focused, routine, rote and superficial: rather than divergent, varied, inventive and deep-seated.

In considering *why* and *how* the Year 11 students in these two case studies learned about fair testing and assessment requirements, the close match between the curricula they experienced in class and their teachers' intended curricula is significant. This finding shows agreement with findings from the literature, that maintain teachers' instructional intentions have a direct bearing on *why* and *how* students learn (e.g., Hargreaves & Fullan, 1992; Lederman, 1999; McGee & Penlington, 2001b; Tytler, 2003). However, given the obvious similarities between the operational curriculum occurring in classrooms and the SiNZC interpretation promoted by the SAS 1.1 it appears the NCEA qualification was a strong influence on the curriculum experienced by Year 11 students in these case studies. This observation is supported by similar overseas experience where high stakes testing and qualifications are also reported to drive classroom practice (e.g., McDonald & Boud, 2003; Orpwood, 2002; Roberts & Gott, 2004, Wiliam, 2000). Qualifications are examples of external determinants of classroom curricula (McGee & Penlington, 2001a) that emanate from 'sites of influence' outside classrooms (English, 1997) and influence the final decisions teachers make about classroom curriculum delivery to their students. The close similarity in each case study between the teachers' intended curricula, the departmental guidelines, and the interpretation of curriculum promoted by the NZQA in its NCEA qualification illustrates the pervasive influence of the NZQA site of influence and the assessment regime underpinning the NCEA qualification on the teachers' planning intentions. This trend is also signalled in recent research by Hipkins (2004) in New Zealand senior science classrooms and suggests that the teachers, and their respective science departments were for the most part acting as conduits for the achievement of that government agency's goals. Many decisions to do with classroom practice were effectively taken out of the individual teachers' hands – instead judgements were made collectively at departmental level. This departmental layer of interpretation was based on guidelines and recommendations from NZQA, which departments and classroom teachers were obligated to follow under school accreditation requirements for NZQA. This departmental layer of interpretation took into account some of the key external determinants of these Year 11 classroom curricula (McGee & Penlington, 2001a), and effectively made decisions that the classroom teachers were obliged to implement in their classroom curricula. These decision included the:

- content of the teaching and learning programme
- manner in which the teaching and learning programmes were to be delivered and assessed, with the emphasis on classroom procedures and arrangements for the practice (formative) and summative assessments and methods of moderation
- timing of the programme delivery
- adoption of the planning template as recommended by NZQA, and the use of exemplar assessment tasks and schedules supplied by the MoE for NCEA as the basis for teaching and assessment materials for use across all classes in the department.

However, as active members of their respective science departments, the teachers in these case studies would have had at least some role in creating this layer of curriculum interpretation and pedagogical and assessment approaches presented in the departmental guidelines. Their participation in meetings concerned with marking and moderation were likely forums for their contributions to be heard and incorporated into the departmental guideline. The strong similarities between the respective departmental guidelines and intended curricula of the teachers in both case studied also lend strong support to the contention that the NCEA qualification had an over-riding influence on the on teachers' instructional intentions.

The actions of individuals from other sites of influence (English, 1997), namely those writers who created the national assessment guidelines and exemplar materials for NCEA and the published text used in the classroom programmes in these case studies, also had a direct effect on students' learning in classrooms because students interacted frequently with these materials in their daily classroom activities and home study.

Conclusions

By examining *what* these students were learning about science investigations, this research found that in both case studies their learning appeared to be focused on a narrow view of scientific inquiry, that is, fair testing, and on mastering assessment techniques. *Why* and *how* this learning occurred stemmed largely from the strong influence the national qualification NCEA, and its interpretation of the science curriculum, was having on decisions affecting the two classroom programmes. This study supports the observations of Black (2001, 2003) that qualification are considered high stakes by schools and teachers and that assessment for qualifications is driving the senior school and classroom programmes in New Zealand. Decisions were made in this study at school and departmental levels, which reflected the importance the two school communities and professional staff placed on their students achieving success in this qualification, and these decisions directly impacted on the content of classroom curricula and the methods teachers used to deliver that content. As a consequence, students were missing out on authentic scientific activity and undoubtedly gaining misleading impressions about the work that scientists do. It would appear that for most NZ students curriculum goals related scientific literacy are not being met

However, in the intervening period since the collection of data for this study NZQA has made some modifications to SAS 1.1 *Carrying out a practical investigation with direction* and introduced more flexibility into the standard and support materials. In October 2005 the standard was re-registered with a number of changes, which seem to introduce more recognition of the complexity of scientific investigation into the standard and give more latitude for teachers to offer students some variety in their approaches to scientific investigation. The revised standard also provides more specific detail about what constitutes 'quality' in a scientific investigation. The achievement criteria are more generic than those in the previous form of the standard, and some former aspects of the accompanying explanatory notes have been given increased emphasis, while some have been dropped and new features introduced. For example:

- greater specificity is provided about what constitutes a directed investigation
- the terms *practical investigation* and *quality practical investigation* are introduced and defined in detail, reflecting the content of the modified achievement criteria. The terms *workable* and *feasible* to describe plans are dropped
- the terms *sample* and *collection of data* are introduced alongside the terms *independent* and *independent variable* respectively in the definition of a practical investigation, and *sampling* and *bias* as possible factors to consider in data gathering in the description of a quality practical investigation. The inclusion of these terms potentially enables students to use approaches to investigation other than fair testing, but because *sampling* and *bias* can have close connotations with fair testing it is possible that fair testing may still prevail in classroom practice unless appropriate exemplary support materials and text are accessible to professional development providers, teachers and students.

- validity of method, reliability of data and science ideas are specified as requirements to consider where relevant when evaluating the investigation

These changes signal more acknowledgement of the nature of scientific inquiry in NCEA assessment procedures for SAS 1.1, and possibly greater opportunity for students to experience authentic scientific investigations and develop higher order thinking skills. This move should give teacher greater autonomy in designing teaching and learning programmes to meet students' learning needs and interests. An overview of exemplary material present on the MoE website for Achievement Standard 1.1 in 2008 revealed one assessment task linked to the new version of the standard. This assessment resource is based on a pattern-seeking investigation. The resource includes a planning and reporting template and assessment schedule similar in format to the fair testing versions, but with terms relevant to pattern seeking and the new requirements of the standard.

Awareness that school-based decisions that focus too much on meeting administrative, logistical and moderation requirements of high stakes qualifications can have detrimental effects on pedagogy and student learning may hopefully prompt schools to re-evaluate the wisdom of these decisions. Finally the views and insights that students have given in this study, about the teaching and learning they experienced and the role they play in these processes, should provide useful information for teachers to reflect on as they evaluate the effectiveness of their teaching and assessment strategies in helping students to achieve quality learning in scientific inquiry.

Deboer (2002) talks of the potential for tension when students are curtailed in their freedom to carry out authentic inquiry by the prescribed content of a standard because teachers and students feel pressured to cover that particular content. The specified nature of the SAS 1.1 and its requirements did contribute to the narrowness of the student experienced curriculum, but it is important to note that the decisions dictating the timing and time allocated to the teaching of this investigative unit were made at school and departmental level, and they were not set requirements of the NCEA qualification or the NZQA. Hipkins (2004) reports a prevailing view among some heads of departments in New Zealand secondary schools that attaining a high number of overall credits in NCEA was superior to gaining excellences but with fewer overall assessment credits. It could be that pragmatic decisions by the schools to provide science courses with high credit numbers, and hence overcrowded curricula, were influenced by similar perceptions of teachers that the quantity of credits gained in NCEA was a criterion by which success in national qualifications could be gauged. The decision to time the unit early in the year, it seems, was one of expediency leaving more time for teachers and students to concentrate on the externally assessed standards later in the year.

Another departmental decision impacting on student learning, that was not required by NCEA and worth noting, concerned the science context in which assessment tasks were set. Choosing a context unfamiliar to students in one case study led to an able student struggling with his explanations of the findings when in an earlier investigation where he was conversant with the background science he had been successful in explaining his findings. This student’s inability to explain results in an unfamiliar science context supports the argument that students need a strong theoretical or conceptual background in the science context of the investigation in order to use scientific theory to make sense of their findings (Harlen, 1998; Leach & Scott, 2003; Luft, 1999).

Summary and Implications

This study sought to find out the nature of the student-experienced curriculum in the New Zealand context as students learn about scientific inquiry for a national qualification from the perspectives of participants in the operational or classroom curriculum. By examining *what* students were learning about science investigations, the research found that the student-experienced curriculum appeared to be focused on a narrow view of scientific inquiry as fair testing, and on acquiring assessment techniques. *Why* and *how* this learning occurred stemmed largely from the strong influence the national qualification NCEA, and its interpretation of the science curriculum, was having on decisions affecting classroom curricula in schools. This study supports the observations of Black (2001, 2003) that qualification are considered high stakes by schools and teachers and that assessment for qualifications is driving the senior school and classroom programmes in New Zealand. Decisions were made in this study at school and departmental level, which reflected the importance school communities and professional staff placed on their students achieving success in this qualification, and this directly impacted on the content of classroom curricula and the methods teachers used to deliver that content.

The qualification interpretation of the science curriculum, in the form of SAS 1.1 and supporting materials, and departmental decisions determining time allocation and timing of the science investigation programme in classes, influenced the didactic pedagogical approaches teachers chose to use, and the strategies used by students to learn. The structure of the qualification, especially the standards-based mode of assessment promoted formative assessment practice with teachers employing many features of convergent formative assessment. However, relatively short teaching and learning programmes before summative decisions were made restricted students' ability to act on formative assessment information to improve their learning. Consequently, student learning tended to focus on procedures and there was little evidence of the higher order thinking skills linked to creativity, evaluating and self-monitoring of learning.

The sway that the qualification interpretation of scientific investigation had on curriculum design and delivery decisions made by schools, departments and classroom lends support to the view that moving from policy document to the operational curriculum in classrooms is not a straightforward process (Atkin & Black, 2003; McGee & Penlington, 2001b), but rather a "cascade of interpreted curricula" (Carr et al. 2001, p. 18). Members of the qualification 'site of influence', for example, in translating the national science curriculum relevant to its purpose of assessment for a qualification, placed emphasis on particular portions of the curriculum that featured fair testing. This practice resulted in a translation contrary to the wider aims of the science curriculum and should alert policy-makers to the importance of conveying a consistent message about student learning outcomes throughout a national curriculum statement. However, introducing more flexibility into SAS 1.1 and support materials should facilitate improved student learning outcomes in terms of authentic scientific inquiry, and give teacher greater autonomy in designing teaching and learning programmes to meet students' learning needs and interests. Awareness that school-based decisions that focus too much on meeting administrative, logistical and moderation requirements of high stakes qualifications can have detrimental effects on pedagogy and student learning, may hopefully prompt schools to re-evaluate the wisdom of these decisions to help students achieve quality learning in scientific inquiry.

References

Adelman, C., Kemmis, S., & Jenkins, D. (1980). Rethinking case study: Notes from the second Cambridge conference. In H. Simons (Ed.), *Towards a science of the singular* (pp. 45-61). East Anglia, UK: Centre for Applied Research in Education, University of East Anglia.

Anderson, L.W., & Krathwohl, D. (Eds.). (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives.* New York: Longman.

Atkin, J.M., & Black, P. (2003). *Inside science education reform: A history of curricular and policy change.* New York: Teachers College Press.

Barnett, J., & Hodson, D. (2001). Pedagogical context knowledge: Toward a fuller understanding of what good science teachers know. *Science Education, 85*(4), 426-453.

Bell, B., Jones, A., & Carr, M. (1995). The development of the recent national New Zealand science curriculum. *Studies in Science Education, 26*, 73-105.

Bell, B. (2005). Pedagogies developed in the Learning in Science Projects and related theses. *International Journal of Science Education, 27*(2), 159-182.

Bell, J. (1999). *Doing your research project*. Berkshire, UK: Open University Press.

Bishop, R. (1997). Interviewing as collaborative storying. *Education Research and Perspectives*, *24*(1), 28-47.

Black, P. (2001). Dreams, strategies and systems: Portraits of assessment past, present and future. *Assessment in Education*, *8*(1), 65-85.

Black, P. (2003). *Report to the Qualifications Development Group Ministry of Education, New Zealand on the proposals for development of the National Certificate of Educational Achievement*. Retrieved March 14, 2003 from http://www.minedu.govt.nz

Bloom, B., Englehart, M., Furst, E., Hill, W., & Krathwohl, D. (1956). Taxonomy of educational objectives: The classification of educational goals. *Handbook I: Cognitive domain*. New York: Longmans Green.

Carr, M., McGee, C., Jones, A., McKinley, E., Bell, B., Barr, H., & Simpson, T. (2001). *The effects of curricula and assessment on pedagogical approaches and on educational outcomes.* Retrieved March 14, 2003, from http://www.minedu.govt.nz

Author (2002). Hamilton, New Zealand: ABA Books.

Crawford, B. A. (2007). Learning to teach science as inquiry in the rough and tumble of practice. *Journal of Research in Science Teaching, 44*(4), 613-642.

Deboer, G. (2002). Student-centred teaching in a standards-based world: Finding a sensible balance. *Science & Education, 11*, 405-417.

Driver, R., Leach, J., Millar, R., & Scott, P. (1996). *Young people's images of science*. Buckingham, UK: Open University Press.

Duggan, S., & Gott, R. (1995). The place of investigations in practical work in the UK National Curriculum for Science. *International Journal of Science Education, 17*(2), 137-147.

Duggan, S., & Gott, R. (2002). What sort of science education do we really need? *International Journal of Science Education*, *24*(7), 661-680.

Duschl, R.A., & Hamilton, R.J. (1998). Conceptual change in science and in the learning of science. In B.J. Fraser & K.G.Tobin (Eds.), *International handbook of science education* (pp. 1047-1065). Dordrecht: Kluwer.

English, E. (1997). The interpretation of the curriculum. Is it all a matter of serendipity? *New Zealand Science Teacher, 85,* 27-33.

English, C., & Wood, R. (1997). The doing of science: A redefinition of procedural knowledge. In B. Bell & R. Baker (Eds.), *Developing the science curriculum in Aotearoa New Zealand* (pp. 67-80). Auckland, New Zealand: Longman.

Erikson, F. (1998). Qualitative research methods for science education. In B.J. Fraser & K.G. Tobin (Eds.), *International handbook of science education* (pp. 1155-1174). Dordrecht: Kluwer.

Garnett, P.J., & Garnett, P. J. (1995). Refocussing the chemistry lab: A case for laboratory-based investigations. *Australian Science Teachers' Journal*, *41*(2), 26-34.

Goetz, J., Preissle.J., & LeCompte, M.D. (1984). *Ethnography and qualitative design in educational research*. Orlando, Fl: Academic Press.

Guba, E., & Lincoln, Y. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage.

Haigh, M. (1995). The Achievement Initiative in Science: Developing Science in the New Zealand Curriculum. In A. Jones, A. Begg, B. Bell, F. Biddulph, M.Carr, M. Carr, J. McChesney, E. McKinley & J. Young Loveridge (Eds.), *SAMEpapers* (pp. 45-74). Hamilton, New Zealand: Centre for Science, Mathematics and Technology Education Research, University of Waikato.

Haigh, M. (2001). Supporting student investigations. *New Zealand Science Teacher, 96*, 39-41.

Haigh, M. (2003). 'Miss, miss, I'm thinking, I'm thinking': Fostering creativity through science education. *New Zealand Science Teacher, 105,* 8-10.

Haigh, M. (2005). Is 'doing science' in New Zealand classrooms an expression of scientific inquiry? *International Journal of Science Education*, *27*(2), 215-226.

Hannay, B., Howison, P., & Sayes, M. (2002). *Year 11 Science. Study guide. NCEA Level 1 edition.* Auckland: ESA Publications.

Hargreaves, A., & Fullan, M.G. (Eds.). (1992). *Understanding teacher development*. London: Cassell.

Harlen, W., & James, M. (1997). Assessment and learning: Differences and relationships between formative and summative assessment. *Assessment in Education, 4*(3), 365-379.

Harlen, W. (1998). The last 10 years: The next 10 years. In R. Sherrington (Ed.), *ASE guide to primary science education* (pp. 23-33). Cheltenham, UK: Stanley Thornes for the Association for Science Education.

Harlen, W. (1999). Purposes and procedures for assessing science process skills. *Assessment in Education*, *6*(1), 129-144.

Harlen, W., & Crick, R.D. (2003). Testing and motivation for learning. *Assessment in Education*, *10*(2), 169-207.

Hipkins, R., Bolstad, R., Baker, R., Jones, A., Barker, M.; Bell, B., Coll, R., Cooper, B., Forret, M., Harlow, A., Taylor, I., France, B., & Haigh, M. (2001). *Curriculum, learning and effective pedagogy: A literature review in science education.* Wellington, New Zealand: Ministry of Education.

Hipkins, R., & Booker, F. (2002). You can't investigate in a vacuum. S*et: Research Information for Teachers 3,* 22-26.

Hipkins, R. (2004, July). *Shifting balances: A study of the impact of assessment for qualifications on classroom teaching.* Paper presented at the annual conference of the Australian Science Education Association, Armidale, Australia.

Hipkins, R., & Barker, M., & Bolstad, R. (2005). Teaching the 'nature of science': Modest adaptations or radical reconceptions? *International Journal of Science Education 27*(2), 243-254.

Hodson, D. (1992). Assessment of practical work: Some considerations in philosophy of science. *Science & Education, 1*, 115-144.

Hodson, D. (1996). Laboratory work as scientific method. *Journal of Curriculum Studies, 28*(2), 115-135.

Hofstein, A., & Lunetta, V. (2003). The laboratory in science education: Foundations for the twenty-first century. *Science Education, 88*(1), 28-54.

Hughes, P. (2004). Mainstream chemical research in school science. *School Science, Review, 85*(312), 71-76.

Author. (2007).. *International Journal of Science Education.*

Hurd, P. (1997). Scientific literacy: New minds for a changing world. *Science Education*, *82*(3), 407-416.

Jenkin, J. (2002). The drive for scientific literacy. *School Science Review, 83*(304), 21-25.

Jenkins, W.J. (1996). The 'nature of science' as a curriculum component. *Journal of Curriculum Studies, 28*(2), 137-150.

Jones, A., & Baker, R. (2005). Curriculum, learning and effective pedagogy in science education for New Zealand: Introduction to special issue. *International Journal of Science Education, 27*(2), 131-144.

Keeves, J.P. (1998). Methods and processes in research in science education. In B. Fraser & K. Tobin. (Eds.), *International handbook of science education* (pp. 1127-1153). Dordrecht: Kluwer.

Knapp, M.S. (2002). *Understanding how policy meets practice: Two takes on local response to a state reform initiative. An occasional paper.* Washington, DC: Center for the Study of Teaching and Policy, University of Washington.

Keiler, L.S., & Woolnough, B.E. (2002). Practical work in school science: The dominance of assessment. *School Science Review, 83*(304), 83-88.

Lange, D. (1988). *Tomorrow's school's: The reform of education administration in New Zealand.* Wellington, New Zealand: Government Printer.

Leach, J., & Scott, P. (2003). Individual and sociocultural views of learning in science education. *Science & Education, 12*, 91-113.

Lederman, N. G. (1999). Teachers' understanding of the nature of science and classroom practice: Factors that facilitate or impede the relationship. *Journal of Research in Science Teaching, 36*(8), 916-929.

Lee, H., & Lee, G. (2001). The National Certificate of Educational Achievement (NCEA): 'Fragile - handle with care'. In I. Livingstone (Ed.), *New Zealand Annual Review of Education* (pp. 5-37). Wellington: School of Education, Victoria University of Wellington.

Luft, J. (1999). Challenging myths. Questioning common myths about science education. *The Science Teacher, 66*(4), 40-43.

Mayer, V., & Kumano, Y. (1999). The role of system science in future school science curricula. *Studies in Science Education*, *34,* 71-91.

McComas, W.F. (1998). The nature of science in science education. Dordrecht: Kluwer.

McDonald, B., & Boud, D. (2003). The impact of self-assessment on achievement: The effects of self-assessment training on performance in external examinations. *Assessment in Education, 10*(2), 209-220.

McGee, C., & Penlington, C. (2001a). *Research on learning, curriculum and teachers' roles. Report 2: Teacher-student interaction.* Hamilton, New Zealand: Waikato Institute for Research in Learning and Curriculum, University of Waikato.

McGee, C., & Penlington, C. (2001b). *Research on learning, curriculum and teachers' roles. Report 3: The classroom curriculum*. Hamilton New Zealand: Waikato Institute for Research in Learning and Curriculum, University of Waikato.

Millar, R., & Osborne, J. (Eds). (1998). Beyond 2000: Science education for the future. London: King's College London.

Ministry of Education. (1993a). *Science in the New Zealand curriculum*. Wellington, New Zealand: Learning Media.

Ministry of Education. (1993b). *Science in the national curriculum.* Wellington, New Zealand: Learning Media.

Ministry of Education. (1993c). *The New Zealand curriculum framework*. Wellington, New Zealand: Government Printer.

Ministry of Education. (1994). *Assessment. Policy to practice.* Wellington: Learning Media.

Ministry of Education. (2002). *Curriculum stocktake report to Minister of Education*. Retrieved September 7, 2003 from http://www.minedu.govt.nz

Ministry of Education. (2003). *New Zealand curriculum project.* Retrieved September 7, 2003 from http://www.tki.org.nz

Moreland, J., Jones, A., & Chambers, M. (2001). Enhancing student learning in technology through enhancing teacher formative interactions. *Set: Research Information for Teachers, 3,*16-19.

Nuthall, G. (1997). Understanding student thinking and learning in the classroom. In B. Biddle, T. Good, & I. Goodson (Eds.), *The international handbook of teachers and teaching* (pp. 1-90). Dortrecht: Kluwer.

Orpwood, G. (2001). The role of assessment in science curriculum reform. *Assessment in Education*, *8*(2), 135-151.

Powell, C.P., & Anderson, R.D. (2002). Changing teachers' practice: Curriculum materials and science education reform in the USA. *Studies in Science Education, 37*, 107-136.

Preece, F.W., & Skinner, N.C. (1999). The National Assessment in Science at Key Stage 3 in England and Wales and its impact on teaching and learning. *Assessment in Education*, *6*(1), 11-25.

Rahm, J., Miller, H. C., Hartley, L., Moore, J. (2003). The value of an emergent notion of apprenticeship: Examples from two student/teacher-scientist partnership programs. *Journal of Research in Science Teaching, 40*(8), 737-756.

Reay, D., & Wiliam, D. (1999). 'I'll be a nothing': Structure, agency and the construction of identity. *British Educational Research Journal*, *25*(3), 343-354.

Reid, N., & Yang, M-J. (2002). Open-ended problem solving in school chemistry: A preliminary investigation. *International Journal of Science Education*, *24*(12), 1313-1332.

Rennie, L. J., Goodrum, M., & Hackling, M. (2001). Science teaching and learning in Australian schools: results of a national study. *Research in Science Education*, *31*(4), 455-498.

Roberts, R., & Gott, R. (2004). Using different types of practical within a problem-solving model of science. *School Science Review*, *85*(312), 113-119.

Ryder, J. (2001). Identifying science understanding for functional scientific literacy. *Studies in Science Education, 36,* 1-44.

Sadler, R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*, 119-144.

Sandoval, W.A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. *Science Education*, *89*(4), 634-656.

Schneider, R.M., Krajcik, J., & Blumenfeld, P. (2005). Enacting reform-based science materials: The range of teacher enactments in reform classrooms. *Journal of Research in Science Teaching*, 42(3), 283-312.

Skamp, K. (2004). *Teaching Primary Science Constructively* (2nd ed.). Victoria, Australia: Thomson.

Shulman, L.S. (1981). Disciplines of inquiry in education: An overview. *Educational Researcher*, *10*(6), 5-12.

Skamp, K. (2004). *Teaching primary science constructively* (2nd ed.). Victoria, Australia: Thomson.

Spindler, G., & Spindler, L. (1992). Cultural process and ethnography: An anthropological perspective. In M.D. LeCompte, W. Millroy, & J. Priessle (Eds.), *The handbook of qualitative research in education.* (pp. 643-680). San Diego, CA: Academic Press.

Spillane, J.P. (2004). *Standards deviation: How schools misunderstand education policy*. Cambridge, MA: Harvard University Press.

Toplis, R., & Cleaves, A. (2006). Science investigations: The views of 14 to 16 year old pupils. *Research in Science and Technological Education, 24*(1), 69-84.

Torrance, H., & Pryor, J. (2001). Developing formative assessment in the classroom: Using action research to explore and modify theory. *British Educational Research Journal*, *27*(5), 615-631.

Tytler, R. (2003). A window for a purpose: Developing a framework for describing effective science teaching and learning. *Research in Science Education, 33*, 273-298.

Wallace, J., & Louden, W. (Eds.). (2002). *Dilemmas of science teaching: Perspectives on problems of practice.* London: Routledge Falmer.

Watson, R., Goldsworthy, A., & Wood-Robinson, V. (1999). What is not fair with investigations? *School Science Review*, *80*(292), 101-106.

Weinburgh, M. (2003). Confronting and changing middle school teachers' perceptions of scientific methodology. *School Science and Mathematics, 103*(5), 222-232.

White, B.Y., & Fredericksen, J.R. (1998). Inquiry, modeling and metacognition: Making science accessible to all students. *Cognition and Instruction*, *16*(1), 3-118.

Wiliam, D. (2000, August). *Integrating formative and summative functions of assessment.* Paper presented to Working Group 10 of the International Congress on Mathematics Education, Makuhari, Tokyo.

Wong, S.L., & Hodson, D. (2009). From the horse's mouth: What scientists say about scientific investigation and scientific knowledge. *Science Education, 93*(1), 109-130.

Table 1. Summary of the teachers' intended learning at *River Valley* and *Mountain High*

Concepts	Skills	Procedural Knowledge
Fair tests Purpose of an investigation as an aim, testable question, hypothesis or prediction Variables - key, dependent and independent Primary and secondary data, qualitative and quantitative data, reliability of data Tables as a systematic format for recording data Graph types (bar and line); graph components such as title, x (independent variable) and y (dependent variable) axes, units and values for axes, plotted points, and lines of best fit Sources of error and systematic errors Equipment names, types and purpose Background/contextual science concepts to the investigation e.g., factors affecting rate of reaction and behaviour of pendulums.	Designing, evaluating, modifying and carrying out a systematic plan for a fair test Determining the purpose of a fair test investigation Identifying, controlling, changing, observing and measuring variables Choosing and using equipment appropriately Determining appropriate range of values for variables Repeating experiments Recording and processing data – tabulating, averaging, graphing Interpreting data, and recognising trends and patterns Discussing findings, linking findings to existing science ideas and drawing conclusions in a written report Evaluating the investigation in the written report (sources of error, improvements).	Knowing how to plan a workable, fair test Knowing that planning requires trialing, evaluating and modifying Knowing why reliable data is needed and how to obtain consistent data Knowing that the findings should be linked to science ideas Knowing how to work as a team Knowing how to interpret the template and assessment schedule requirements of tasks for the internal Science A.S. 1.1 at achievement, merit, and excellence levels.
*At River Valley Jenny added 'Good science' (the science that real scientists do), and 'school science' (the portrayal or simulation of science experienced by students in school); systematic errors; and the concept of controls *At Mountain High Kathy provided an experimental plan which included an aim, list of equipment and an experimental method, a format for scientific reports and coverage of the relationship between two quantities when change in one causes change in the other	*At River Valley Jenny also included some trialing of plans	*At River Valley Jenny also dealt with how to recognise and account for errors in measurement; and recognising that the planning and carrying out of investigations required for Science A.S. 1.1 more closely resembles 'good science', than most 'school science' *At Mountain High Kathy added knowing that the findings should be linked to the science behind the investigation; and knowing when assumptions can be made and the limitations of those assumptions

Table 2. Summary of the key influences on *Why* and *How* students learned

The Content of the Teachers' Intended Curricula	The Pedagogical Approaches, Strategies and Capabilities of their Teachers	The Learning Strategies that Students Employed
Teachers delivered content in the teaching and learning programmes specifically targeted at fair testing and the assessment requirements of SAS 1.1 Teachers' decisions about lesson content were governed by their respective school departmental guidelines for delivering the SAS 1.1 - all teachers in the departments were obliged to follow these guidelines. Departmental guidelines were similar in each school since each school looked to materials provided by government agencies to support learning programmes for the SAS 1.1 i.e., planning templates, and exemplar assessment tasks and schedules *The exposure of students at *River Valley* to the notions of 'good science' as opposed to 'school science' in their learning, probably stemmed from their teacher's own knowledge base and beliefs about the nature of scientific investigation and her personal experience of scientific research.	Departmental guidelines produced many commonalities in the pedagogical strategies teachers employed - they effectively decided the: *manner* in which the teaching and learning programmes were to be delivered and assessed; *timing* of the programme delivery and; *adoption* of the planning template and exemplar assessment tasks and schedules. As a result teachers' pedagogical approaches were predominantly didactic in nature Students identified particular common teaching strategies that helped their learning, including: provision of the opportunity to do practice investigations and write-ups for assessments in groups: direct instruction from knowledgeable teachers; provision of a planning template and assessment schedules and; feedback they received from teachers and fellow students after assessments. Convergent formative assessment practice underpinned *why* and *how* students were succeeding in many aspects of their learning. Explicit sharing of learning goals, success criteria and learning progress with students was achieved via the use of exemplars.. The timing of the teaching and assessment early in the school year appeared to limit students' opportunities to consolidate and improve their learning in a wide range of contexts, and to develop the tacit, intuitive knowledge required for effective investigating in science. The teaching decision to set both the formative and summative investigations in the same familiar science context possibly gave students at *Mountain View* the opportunity to make meaningful links with their new experiences more readily than students at *River Valley* , where the background science in the summative assessment was unfamiliar to students and they had had little exposure to the phenomenon being investigated	Students often played a mediating role in their learning, at times consciously choosing when and how to engage from a range of personally preferred learning strategies. Learning choices were often related to perceptions students had about what was valuable or important to learn and who was best suited to assist their learning at given times, and feelings of self-esteem and self-confidence: - NCEA was an important personal goal for most students, and they were prepared to learn what was required of them in order to demonstrate achievement of the standard at particular levels of attainment - high value was placed on being able to work and collaborate with peers – students appreciated the convenience and ease of sharing knowledge and expertise to problem solve, and to clarify misconceptions and/or confirm understanding in the relatively safe forum of pairs/small groups of students. They realised some interactions between peers could also be detrimental to learning, and lack of effective teamwork was seen to compromise intended learning work on at least one occasion. -students were ambivalent about the value of peer assessment in promoting and facilitating their learning, generally because they questioned the credibility and capability of their peers to assess as accurately as their teachers. While it was difficult to judge individual students' capabilities on the basis of negotiated group plans, the collaborative planning process tended to give more group members the potential to secure relevant and reliable data, and in turn the chance to process and interpret data, draw conclusions and evaluate their findings.