1-1-1998

Science laboratory investigations in year 9: Instruction and assessment

Pamela J. Garnett
Edith Cowan University

SCIENCE LABORATORY INVESTIGATIONS IN YEAR 9: INSTRUCTION AND ASSESSMENT

Pamela J Garnett

B Sc, Dip Ed, Grad Dip Sci Ed, M App Sci

This Thesis is presented for the degree of Doctor of Philosophy at

EDITH COWAN UNIVERSITY

PERTH WESTERN AUSTRALIA

1998

USE OF THESIS

I certify that this thesis does not, to the best of my knowledge and belief:

(i) incorporate without acknowledgment any material previously submitted for a degree or diploma in any institution of higher eduction;

(ii) contain any material previously published or written by another person except where due reference is made in the text; or

(iii) contain any defamatory material.

Pamela Joy Garnett

ABSTRACT

The purpose of this research was to develop, implement and evaluate science investigation work based on a cognitive apprenticeship model of instruction and linked to different assessment procedures. Data were gathered to evaluate the effects of the instruction and approaches to assessment on teachers and students, and on the development of students' investigation competency. Year 9 students performed six science investigations in which they examined a relationship between variables. Groups of three students worked together to choose the independent variable, plan their investigation, collect and analyse their data, and to evaluate their findings. The cognitive apprenticeship model of instruction included teacher modelling, coaching, scaffolding and fading, articulating and involving students in self-reflective and metacognitive practices. The three classes which participated in the study (n = 66) experienced different assessment regimes, teacher assessed and norm referenced, teacher assessed and criterion referenced and student assessed and criterion referenced assessment.

The study was a naturalistic inquiry and data were collected from numerous sources including a pre and posttest pencil and paper Test of Science Investigation Skills, pre and posttest investigations in which students were assessed from their written responses on an Investigation Planning and Report Sheet, student questionnaires, teacher and student group interviews, and audio and video data. The qualitative data were summarised and interpreted as 76 assertions relating to the themes of the research; investigation competencies, the cognitive apprenticeship model of instruction and the assessment regimes. Sixteen general assertions which were considered to be more general research findings were then formulated.

Students' investigation competencies improved significantly as determined by their pre and posttest performances on the Test of Science Investigation Skills, and by their performances on the Investigation Planning and Report Sheet. In addition to improving specific investigation competencies such as planning and conducting investigations, processing data and evaluating investigations, students also perceived that they improved social and workplace skills including working cooperatively, attending to detail, managing time and being organised. Difficulties students experienced in performing investigations were also identified.

Based on improved students' performance, the cognitive apprenticeship model of instruction could be deemed as effective in teaching and learning science investigation competencies. As implemented in the study, weaknesses in the implementation of teacher modelling were exposed and it was suggested that modelling would be more beneficial if it were to occur in response to requests from students rather than at the behest of the teacher as an introduction to a learning activity.

Classes experiencing teacher assessment, both norm referenced and criterion referenced, achieved similar gains in investigation competencies and the feedback to students following these assessments was similar in quality and quantity. Students in the student assessed criterion referenced class made more modest improvements in *Planning investigations* and *Conducting investigations*. These students lacked opportunities for high quality teacher feedback. Clearly these data need to be interpreted with caution because other factors in the learning milieu were not controlled.

The implications arising from the research for classroom practice, addressed factors contributing to science investigation competency, the cognitive apprenticeship model of instruction and the assessment of science investigations.

ACKNOWLEDGMENTS

This research has been achieved with the support and guidance of many friends. My supervisor, Mark Hackling, has encouraged me and provided me with thoughtful, well-considered advice and feedback. His openness to discussion and willingness to help have shaped the final outcome. Leonie Rennie, my co-supervisor, has provided constructive criticism and valuable advice.

To my own science staff I owe a huge debt of gratitude. They were helpful and supportive at all times. I am most appreciative of their cooperation and their acceptance of the intrusive audio and video equipment in their classrooms. I am also thankful to my school for granting me leave to write-up my thesis, and more importantly, for providing a teaching environment that encourages and supports professional growth and excellence in teaching and learning.

To my husband, Patrick and daughters Sally, Megan and Katherine thank you for your love and encouragement. Special thanks go to Katherine for her computer diagrams (Figures 1, 2 & 22).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES (continued)

LIST OF FIGURES

LIST OF FIGURES (continued)

CHAPTER 1

INTRODUCTION

Science Investigations as a Domain of Study

The role of practical laboratory work in science has been questioned by numerous researchers in science education (Hodson, 1990; Lynch, 1987; Woolnough & Allsop, 1985). Woolnough and Allsop claim that one reason for the failure of many science courses is the attempt to use practical laboratory work for aims to which it is ill-suited, namely the teaching of theoretical concepts, instead of developing process and problem solving skills, and developing a 'feel' for natural phenomena. According to Lloyd (1992), a structured or cookbook approach is still the overwhelming choice in laboratory manuals. Bryce (1994) reports that in Western Australian schools teachers determine the problem to be investigated, the apparatus to be used and the procedure to be followed in 84% of laboratory activities. Clearly these activities provide few opportunities for students to formulate hypotheses and to design experiments or procedures. They are also characterised by insufficient discussion concerning the limitations of the methodology and underlying assumptions and the degree of confidence that can be placed in the data. Therefore it is not surprising that for many students "a 'lab' means manipulating equipment but not manipulating ideas" (Lunetta, 1998, p. 250), and that in Western Australia research (Hackling & Garnett, 1993) indicates that secondary students have poorly developed skills of problem analysis, planning and carrying out controlled experiments, basing conclusions on obtained data, and recognising the limitations in the methodologies of their investigations.

The United Kingdom has embraced the notion of science investigations with more enthusiasm than other countries. With the inception of the National Curriculum

in the United Kingdom, an attainment target focusing on experimental and investigative science has been identified which includes planning experimental procedures, obtaining evidence, analysing evidence and drawing conclusions, and evaluating investigations. In Australia, the Working Scientifically strand of the national science curriculum framework (Australian Education Council, 1994a) and accompanying profile of learning outcome statements (Australian Education Council, 1994b) reflects a similar emphasis on science investigations. Similarly, in Western Australia investigating is identified as one of nine essential learning outcomes in the draft Science Learning Area Statement (Curriculum Council of Western Australia, 1997). In the United States of America experimental and investigative science is referred to in the National Science Education Standards as "scientific inquiry" (National Academy of Sciences & National Research Council, 1996, p. 23).

Problem

In Western Australia, considerable interest has been generated in science laboratory investigations and the challenge is to improve students' attainment of investigation skills. The problem is to determine how best to achieve this goal. For investigative work to translate into successful classroom practice there is a need to integrate complementary theories of learning with instructional models and assessment procedures. This study will evaluate the effect of a cognitive apprenticeship model of instruction combined with various assessment procedures on students' attainment of science investigation skills.

Rationale and Significance

The national curriculum/standards movements in the United Kingdom, Australia and the United States of America has provided teachers with a framework to refocus approaches to instruction and assessment in the laboratory. Traditional

closed laboratory exercises which comprise verifying a stated principle or relationship or seeking patterns or relationships in data (Lunetta, 1998) need to be used in conjunction with open, problem solving investigations to give students opportunities to develop science investigation skills. The cognitive apprenticeship model of instruction was selected because learning to conduct investigations has been likened to learning a craft (Millar, 1991), and because a key aspect of instruction is the formation of a conceptual model of the task (Collins, Brown & Newman, 1989), a notion that is compatible with the holistic learning of investigation competencies advocated by Hodson (1992). Different assessment regimes were selected for the classes participating in the study because currently norm referenced assessment procedures are being questionned and challenged by criterion and standards referenced assessments. The study seeks to bring together theories of learning and models of instruction that are appropriate for the development of science investigation skills and complement these with appropriate assessment procedures. Research is needed to develop, implement and evaluate the new science laboratory curriculum in terms of its impact on the roles and responses of teachers and students, and on students' attainment of science investigation competencies.

This study will make an original contribution to the literature on instruction, assessment and learning from science investigations, and a direct contribution to laboratory curriculum and the practice of laboratory instruction.

Purpose and Research Questions

The purposes of this study are to develop, implement and evaluate Year 9 science laboratory investigation work based on a cognitive apprenticeship model of instruction and linked to three different assessment procedures. Data are to be gathered to evaluate the effects of the instruction and approaches to assessment on

teachers and students, and on the development of students' investigation competence. More specifically the research addresses the following questions.

1. What science investigation competencies and understandings are developed by students during the instructional program implemented in the study and what difficulties do students experience?

2. In the teaching and learning of science investigation competencies how effective is the cognitive apprenticeship model of instruction?

3. What effect do different assessment procedures including teacher assessed norm referenced, teacher assessed criterion referenced and student assessed criterion referenced assessments have on students' learning of investigation competencies?

Definition of Terms

The teaching strategies defined in the cognitive apprenticeship model of instruction vary slightly in intent, meaning and application, depending on their source. In this research the strategies are defined as follows.

Teacher modelling. This involves a teacher performing a science investigation with students observing the investigation. The teacher demonstrates the procedures to be learned and models his/her decision-making processes.

Teacher guidance. This involves the following strategies:

(a) **Coaching:** This is an impromptu response by the teacher to help students perform an investigation skill or operation. It involves

discourse between the teacher and student and hence includes teacher-student interactions or expert-novice articulation.

(b) **Scaffolding:** This is a predetermined strategy or structure that is used to facilitate, structure and support learning and is based on the teacher's conceptual model of the task and expectations of students' difficulties.

(c) **Fading:** This is the gradual withdrawal of modelling, coaching and scaffolding as students become increasingly competent.

Articulating. This is the discourse between students and the teacher, and between students as they perform the investigations.

Self-reflective and metacognitive skills are thinking skills which involve learners reflecting on their learning.

Investigation competencies include skills, competencies and understandings that are required for students to perform investigations. Investigation competency also includes personal qualities and attributes needed by students to be able to work in a group to complete . investigation.

An **open investigation** refers to one in which the student can choose a course of action. For the investigations in this research the problem is defined for the students, however, they are required to choose the independent variable to investigate and their method of investigation.

Overview of the Thesis

An instructional program, grounded on the cognitive apprenticeship model of instruction, was designed to develop students' science investigation competencies.

Three classes of approximately 22 students participated in the program of 30 lessons. The classes had different assessment regimes, teacher assessed and norm referenced, teacher assessed and criterion referenced, and student assessed and criterion referenced assessment. Following the literature review in Chapter 2, the methodology is outlined in Chapter 3 with a description of the instructional program and the assessment regimes.

In order to answer the research questions a variety of data gathering procedures was used. Pre and posttest data, presented in Chapter 4, were gathered to gauge the impact of the instruction and assessment on students' investigation competencies. Students' responses to the instructional program and the assessment regimes were accessed through questionnaires (Chapter 5) conducted with the whole cohort, and interviews conducted with three groups of three students (Chapter 7). Teachers' responses to the instruction and assessment were gathered through interviews and are presented in Chapter 6. Audio and video data of a group of three students from each of the classes, were used in conjunction with portfolios of students' work to gain insights into the learning milieux in which the program operated. These data are presented in Chapter 8.

Throughout Chapters 4 to 8 single event assertions are derived from different data sources. Chapter 9 clusters the assertions into themes; investigative competencies, the cognitive apprenticeship model of instruction and assessment. It also triangulates the assertions by indicating those that are supported by more than one data source. Research findings or more general assertions are postulated from the single event assertions and are discussed in relation to the research questions (Chapter 10). In conclusion, Chapter 10 also presents a brief summary of the research, the implications for teaching and the theoretical framework, and the limitations of the research. In addition, it makes recommendations for future research, and discusses the contribution of the research to science education.

CHAPTER 2

LITERATURE REVIEW

Overview of the Chapter

Salient factors that impact on this research are indicated in Figure 1. These include learning theory, the cognitive apprenticeship model of instruction, the nature of science investigations and assessment issues associated with the investigations. The cognitive apprenticeship model of instruction includes modelling, coaching, scaffolding, fading, articulating, and self-reflective and metacognitive skills. Several studies that utilise these teaching strategies are discussed. Research literature associated with the nature and purpose of science investigations, and investigation competencies is discussed. Issues are raised such as the holistic development of investigation competencies, the linking of investigation competencies with conceptual understandings, and factors affecting students' abilities to perform investigations. Group work is also addressed because investigations are normally conducted in small collaborative groups. Finally, aspects of assessment are discussed and, in particular, the assessment of science investigations is addressed.

Learning Theory

Recently the learning of science has been influenced by the constructivist learning theory (von Glaserfeld, 1989) which proposes that students actively construct new meaning by using their present conceptual frameworks to interpret new information in ways that make sense to them. Driver, Asoko, Leach, Mortimer and Scott (1994) have described how our understanding of students' learning has

been informed by the debate between the personal and social constructivist traditions and a consideration of the nature of scientific knowledge. The personal constructivist tradition (Cary, 1985) emphasises the learner's personal construction of knowledge and the concepts that individuals develop about natural phenomena. Learning is viewed as conceptual change. In comparison, the social constructivist tradition (Edwards & Mercer, 1987) recognises science as symbolic, and socially constructed and communicated. The social context in which learning takes place is crucially important if individuals are to construct scientific interpretations of the world around them.

All cognition is situated in the context of the activity associated with the learning, and the activity "is not separable from or ancillary to learning and cognition. Nor is it neutral. Rather it is an integral part of what is learned" (Brown, Collins & Duguid, 1989, p. 32). This notion of situated cognition challenges teaching practices that implicitly assume that conceptual knowledge can be abstracted from situations in which it is learned and used in other contexts. By using the teaching of reading, writing and mathematics as examples, Collins, Brown and Newman (1989) argue that cognitive apprenticeship methods enculture students into authentic practices through activity and social interaction. Hence, the cognitive apprenticeship model of instruction is closely aligned with the social constructivist tradition.

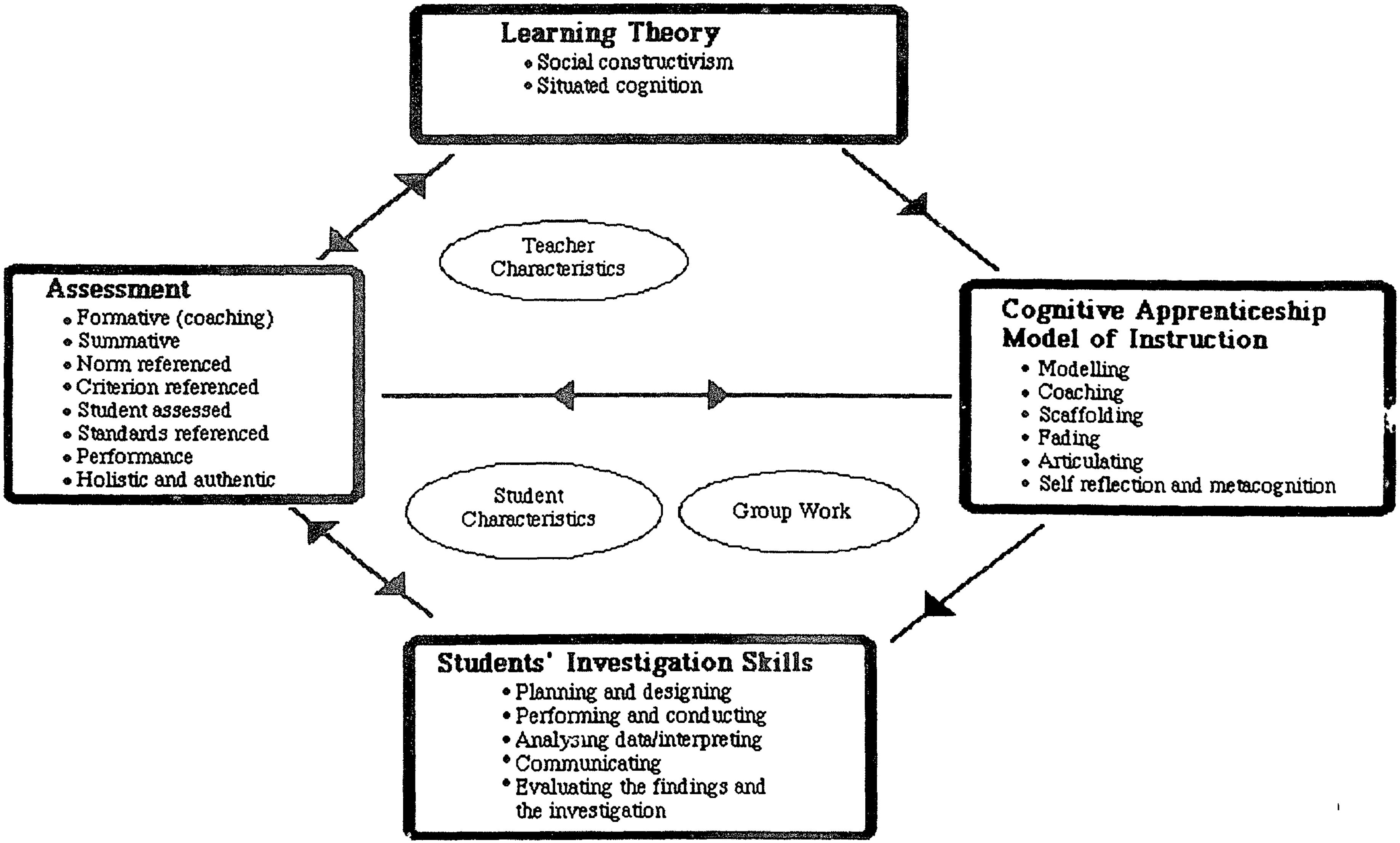


Figure 1. Factors impacting on the research

The Cognitive Apprenticeship Model of Instruction

The traditional craft apprenticeship focuses closely on learning specific methods for carrying out tasks in a specific domain (Collins et al. 1989). The parallel with this study is students' acquisition of competencies and methods for carrying out science investigations, which according to Millar (1991) are more like a craft than the following of rules or a hierarchy of processes. In the traditional apprenticeship, Collins et al. (p. 456) state "the apprentice repeatedly observes the master executing (or modelling) the target process before he or she attempts to execute the process with guidance." Lave (as cited in Collins et al.) contends that apprentices learn the processes through a combination of what she calls observation, coaching, and practice, which from the teachers' perspective may be called modelling, coaching and fading. The apprentice repeatedly observes the master executing or modelling the target process. The apprentice then attempts to execute the process with guidance. A key aspect of guidance or coaching is the provision of scaffolding, which is the support, in the form of reminders and help, that the apprentice requires to approximate the execution of the entire composite of skills comprising the whole activity. Collins et al. contend that observation plays a key role. Lave hypothesises that observations aid learners in developing a conceptual model of the task or process, prior to attempting to execute it and that the provision of a conceptual model is important for three related reasons. First, it provides an advanced organiser to allow learners to concentrate more of their attention on the execution of the process than would otherwise be possible. Second, a conceptual model provides an interpretive structure for making sense of the feedback, hints and corrections during interactive coaching sessions. Third, it provides an internalised guide for independent practice.

"Cognitive apprenticeship methods try to enculturate students into authentic practices through activity and social interaction in a way similar to that evident - and

evidently successful - in craft apprenticeship" (Brown et al. 1989, p. 37). Brown et al. add that the cognitive apprenticeship model supports learning in a domain by enabling students to acquire, develop and use cognitive tools in an authentic activity where the activity is central to learning. The notion of cognitive apprenticeship implies that the skills acquired are well beyond physical skills usually associated with craft apprenticeships and that the kinds of cognitive skills normally associated with conventional schooling are addressed.

Collins et al. (1989, p. 457) contend that the transition from the traditional craft apprenticeship to the cognitive apprenticeship has two benefits for learning. First, the method is aimed primarily at teaching the processes that experts use to handle complex tasks and as a result "conceptual and factual knowledge are exemplified and situated in the context of their use." They add that the dual focus on expert processes and situated learning help solve the educational problems of brittle skills and inert knowledge. Second, the focus is on learning cognitive and metacognitive skills and processes rather than physical skills and processes.

Two major differences (Collins et al., 1989) exist between the traditional apprenticeship and the cognitive apprenticeship. First, the tasks given to apprentice workers arise from the demands of the workplace, while at school the tasks given to learners arise from pedagogical concerns. Second, with schooling there is necessarily a need to decontextualise knowledge and skills so that they can be used in different settings. Therefore school learning needs to be situated in diverse settings so that students learn how to apply their skills in varied contexts. In addition, the application of knowledge and skills in different settings needs to be articulated as fully as possible by the teacher whenever opportunities arise.

Researchers have defined slightly different teaching strategies to comprise the cognitive apprenticeship model of instruction and because of this the

interpretation of terms as applied to this study has been presented in Chapter 1 (p. 4). In a seminal paper Collins et al. (1989) identified six teaching methods that characterise the instructional model; modelling, coaching, scaffolding, articulation, reflection and exploration. More recently, Hennessy (1993) described the features based on the work of Collins et al. as modelling, coaching, scaffolding, fading, articulating and encouraging learners to reflect on their own problem solving strategies through self-reflection and metacognition. Then, she continued to discuss articulation, modelling, fading and scaffolding.

In these articles (Collins et al., 1989; Hennessy, 1993) and those of other researchers (Javela, 1996; Roth, 1995) the definitions of and distinctions between the teaching strategies in the model become somewhat blurred. **Modelling**, however, is consistently defined as involving "an expert carrying out a task so that students can observe and build a conceptual model of the processes that are required to accomplish the task" (Collins et al., 1989). According to Hennessy this involves making tacit knowledge explicit.

Coaching is defined by Collins et al. to consist of "observing students while they carry out a task and offering hints, scaffolding, feedback, modeling, reminders, and new tasks aimed at bringing their performance closer to expert performance" (p. 481). Hence coaching includes scaffolding and modelling. They add that coaching may be used to direct a student's attention to a previously unnoticed aspect of the task or to remind them of something that they have overlooked. The notion of coaching is consistent with the purpose of formative assessment which is described as "helping pupils during the process of learning" (Radnor & Shaw, 1995, p. 132).

Scaffolding refers to "the support the teacher provides to help the student carry out a task" (Collins et al., 1989, p. 482) and because providing support may be construed as providing help (coaching), again the teaching strategies converge.

According to Collins et al., central to scaffolding is the identification of the student's current level of skill and the provision of intermediate tasks that lead to the target activity. Similarly, Vygotsky (1986, p. 189) has argued that it is important to identify not just what students have achieved but what it is that students *might* achieve and work to expand the limits of the learner's confidence and competence in the "zone of proximal development". It is important that learning is facilitated by "scaffolds" (Bruner, 1985, p. 28) and students are set tasks within the zone of proximal development and provided with appropriate support. Similarly, Hennessy (1993) recognises the contribution of scaffolding to learning and says that "subsequent interpretations and applications of the notion of apprenticeship have without exception focused on the tutor's implicit theory of the learner as being a crucial element of the scaffolding process" (p. 12). The notion of scaffolding has been implemented as the provision of suggestions and help in Palinscar and Brown's (1984) reciprocal teaching, and as cue cards in Scardamalia and Bereiter's (1985) procedural facilitation of writing. According to Hennessy (p. 13) classroom teachers carry out an "on-line diagnosis" of students' responses (similar to formative assessment in the preceding paragraph) and that from such diagnoses scaffolding follows. She adds that the essential features of scaffolding in the classroom setting have not yet been identified. Solomon's (1988, p. 104) statement, "we know that the pupils will not find out about scien ´ic models by unaided discovery learning, because observation is most forcefully guided by beliefs already held," implies that it is necessary to provide directions for students' learning, particularly when existing beliefs are to be challenged. This view is supported by Hodson (1996).

Roth (1995) places a slightly different interpretation on the teaching strategies and merges coaching, articulating and scaffolding. He links coaching and articulating as follows, "coaching is a more inclusive notion to describe teacher-student interactions" (p. 243). He also links coaching and scaffolding as he adds that "a coach also designs the tasks in such a way that students can practise their

knowledge and skills in settings that are challenging but at a complexity appropriate to the student's current abilities" (p. 243).

The contribution of practice to learning is not explicitly discussed in the studies mentioned in this literature review although the provision of repeated opportunities to practise decision-making on complex tasks is implied. Coaching and scaffolding, followed by the withdrawal of support suggest that practice is an element of the instructional model. The gradual reduction of levels of support given to students is referred to as **fading** (Collins et al., 1989; Hennessy, 1993; Roth, 1995).

Articulation includes "any method of getting students to articulate their knowledge, reasoning, or problem solving process in a domain" (Collins et al. 1989, p. 482). Collins et al. document three methods of articulation; questioning students to lead them to a particular conclusion, encouraging students to express their thoughts as they carry out problem solving, and having students assume the critic and monitor role in cooperative activities. The instructional model does not explicitly address articulating that occurs amongst students but focuses more on expert-novice verbal interactions that are more consistent with the notion of apprenticeship. According to Vygotsky (1986) and Hennessy (1993) articulation is encouraged through providing students with alternative viewpoints and counter examples. This implies that the student's perspective is compared with that of an expert, the teacher, and that verbal interaction results in the resolution of differences or the strengthening of common views. Hennessy adds that "this process makes normally hidden mental processes overt" and the "aim is to give learners control over their own learning processes and the confidence to engage in critical analysis" (p. 12).

"**Reflection** enables students to compare their problem solving processes with those of an expert, another student, and ultimately, an internal model of expertise" (Collins et al., 1989, p. 482). This may be achieved through 'replaying' the performances of the expert and novice for comparison, and/or by a post mortem of a student's own problem solving process. The development of these skills is consistent with the metacognitive skills that Australian students were encouraged to develop in the Peel Project (Baird, 1986a; Baird, 1986b). In this project students were trained to practise applying evaluative cognitive strategies during lessons. The findings of the Peel Project indicate that students became more informed, purposeful learners, exerting greater control over their learning.

Exploration, the final strategy in the cognitive apprenticeship model, involves pushing students into a mode of problem solving on their own (Collins, et al., 1989). It involves the fading of supports and problem setting so that students are given opportunities to learn how to frame questions and problems that they are interested in researching.

Numerous researchers have used aspects of the cognitive apprenticeship model of instructions in classroom studies. In the reciprocal teaching of reading comprehension, Palinscar and Brown (1984) based the instruction of two groups of seven students as well as individual students, on modelling and coaching in four strategic reading comprehension skills; formulating questions, summarising, clarifying and predicting. Students were engaged in a set of activities that helped them form a new conceptual model of the reading task. They practised the reading strategies and metacognitive skills necessary for expert reading. The teacher modelled expert strategies in a context, and shared the strategies directly and immediately with the students. Scaffolding was provided for the students when they were at an impasse. These techniques proved "remarkably effective in raising students' scores on reading comprehension tests" (Collins et al. 1989, p. 460).

In the teaching of writing, Scardamalia and Bereiter (1985) used a combination of modelling, coaching, scaffolding and fading to give students a grasp of the complex activities involved in writing. According to Collins et al. (1989) students were able to build a new model of the writing process that more closely emulated that of expert writers, and they developed an increased awareness of the self-reflective and metacognitive processes involved in expert writing.

To teach mathematical problem solving, Schoenfeld (1985) formulated a set of heuristic strategies based on the problem-solving strategies of Polya (1945). He used aspects of the cognitive apprenticeship instructional model including modelling, coaching, scaffolding, fading and encouraging students to reflect on their practice in a variety of activities to teach students the heuristic strategies. In addition, he used a post mortem analysis that is described by Collins and Brown (1988) as an abstracted replay involving the recapitulation of critical decisions and actions. Schoenfeld's method consisted not only of applying problem solving strategies, but also reasoning about managing problem solving using heuristics, control strategies associated with making decisions, and beliefs about one's own ability. One of his aims was to enculture students into the world through a mathematician's eyes and to provide them with the tools of a mathematician.

More recent research (Chee, 1995; Ertmer & Cennamo, 1995; Jarvela, 1996; Johnson & Fischbach, 1992; Pieters & DeBruijn, 1992; Roth & Bowen, 1995; Volet, 1991) has also focused on aspects of the cognitive apprenticeship instructional model. For example, Ertmer and Cennamo used the instructional model to foster the development of 'expert' designers in an instructional design course that they implemented. The instructional design course involved the following; 'think alouds' in which the teachers analysed contexts unfamiliar to the students (modelling); teachers meeting with students who sought clarification of design ideas (coaching);

students comparing their ideas with those from other sources (reflection); students verbalising their thought processes as they carried out specific design tasks (articulating); and students broadening their perspective through role plays and participation in case studies (exploration). Ertmer et al. reported that "there seems to be much less re-doing, and much more understanding of what makes a good design" (p. 56). They added that they were successful in encouraging students to become more reflective in their thoughts, reactions and decisions. Problems they identified included a perception that they had sacrificed some "breadth in terms of content coverage", that it was difficult to select suitable design problems, and that the design project took a long time to assess.

Jarvela (1996) chose the cognitive apprenticeship model to organise instruction in a technologically rich learning environment in which 13 to 14 year old boys investigated and modelled the control technology principles of an automatic washing machine using Lego Logo ™. She analysed, in terms of scaffolding, modelling and reflection, video recordings of four pairs of students working for nine hours. Jarvela associated scaffolding and coaching with articulating. She reported that teachers either scaffolded interactions with students or coached them, and that this was a function of the student's ability, the phase of the student's problem solving, and their motivation. She defined scaffolding as "giving small hints and stimulating statements for students' strategic approach," and coaching as "giving explicit advice or concrete cues" to improve a student's involvement in the task (p. 98). Jarvela distinguished between teacher "global modeling in front of the class" and teacher "situation-specific modeling" with a small group of students (p. 100). She claims that situation specific modelling was experienced "more reciprocally" and was "directed to the actual problem that the students had" (p. 101). She contends that situation-specific modelling has the potential to promote spontaneous, more advanced exploratory activities among the students. Jarvela does

not clearly define 'reflection' and used questions that students and teachers asked as evidence for spontaneous reflection.

In a Grade 8 science classroom, Roth and Bowen (1995) analysed the processes of knowing and interacting in an open-inquiry learning environment in which teachers used a cognitive apprenticeship metaphor. Students were challenged to work "like a real biologist" (p. 84) and possible behaviours and thoughts of biologists were described (modelled) at the start of the open-inquiry in which they were instructed to find out all they could about the biotic and abiotic factors of a 35 m^2 ecozone on the school site. From the cognitive apprenticeship perspective Roth and Bowen (p. 91) report that "students constructed increasingly complex and interesting research problems." Support for the students faded as they became more proficient with their inquiry. It is inferred that teacher guidance was in the form of carefully constructed teacher questions to provoke problem solving. Roth and Bowen (p. 91) said that "despite careful construction, teacher 'problems' were far from unequivocal, so that students constructed unintended meanings that led them sometimes to solve problems different from those intended." They contend that students did this in three ways. First, if there were no specific "correct" answer to the teacher's question then the students appropriated the problem and solved it as their own problem. Second, students constructed their own version of the problem for which they perceived a "correct" solution to exist and this often led to data and results that would differ from experts. Third, they constructed elaborate contexts to frame the problem so that it made sense to them. The researchers also reported that students' private knowledge was distinct from their public knowledge and that students "(a) constructed private understandings continuously, interpreted what they saw fit into their current understanding and raised questions when they did not understand what they measured or observed; (b) collaborated to achieve mutual understandings of problematic situations; and (c) interacted frequently with other students in order to get the work done" (p. 110).

Each of the preceding studies that documented specific aspects of the cognitive apprenticeship model of instruction, reported improved student achievement of the target attribute; reading comprehension (Palinscar & Brown, 1984); writing (Scardamalia & Bereiter, 1985); mathematics problem solving (Schoenfeld, 1985); instructional design processes (Ertmer & Cennamo, 1995); design technology processes (Javela, 1996) and open science inquiry (Roth & Bowen, 1995). They had the common objective of providing learners with strategies to lead them from novice through competent and towards expert understandings and processes.

Laboratory Work in Science

During the 1980s science laboratory practices in Australian schools contributed little to students' learning. Gallagher and Tobin (1987) and Tobin and Gallagher (1987) described students gathering data without comprehending the meaning of their actions and without reflective thought because they reduced the cognitive demands of the tasks to a minimal level. Students went about their laboratory work in a leisurely atmosphere and spent much of their time off-task, socialising with their peers. Despite this Tobin (1990, p. 403) states, "Laboratory activities promise so much in term of students being able to solve problems and construct relevant science knowledge."

According to Lunetta (1998), a predominant pattern in the teaching of science has been "telling the story of science" (p. 251) and that compatible with this pattern is laboratory work that "engaged students principally in following ritualistic procedures to verify the story that had been told" (p. 251). The view of laboratory work espoused by Lunetta is consistent with other researchers. Lloyd (1992) claims

that a structured or cookbook approach overwhelmingly dominates laboratory manuals and Bryce (1994) reports that in Western Australian schools teachers determine the problem to be investigated, the apparatus to be used and the procedure to be followed in 84% of laboratory activities. The challenge for science educators and teachers is to move students beyond the mere following of laboratory procedures to engaging them intellectually with meaningful laboratory experiences in which they "construct shared understanding of scientific concepts in a community of learners in their classroom" (Lunetta, 1998, p. 252).

Recently in Australia, the United Kingdom and the United States of America, national science curriculum, standards and assessment documents (Australian Education Council, 1994a, 1994b; Curriculum Council of Western Australia, 1997; National Academy of Science & National Research Council, 1996; United Kingdom Department of Education and Science, 1988) have been produced and refocus attention on the role of science laboratory work. The purposes of laboratory work have been articulated by numerous writers (Boud, Dunn & Hegarty-Hazel, 1986; Garnett, Garnett & Hackling, 1995; Garnett & O'Loughlin, 1989; Hegarty-Hazel, 1990; Hodson, 1988; Lunetta, Hofstein & Giddings, 1981; Solomon, 1988; Woolnough, 1991) and may be considered to develop conceptual understandings, techniques and manipulative skills, investigation skills and competencies, and affective outcomes. Hence, if the intent of a laboratory activity is to achieve a particular outcome then the suitability of the activity needs to be considered in relation to the desired outcome (Duggan & Gott, 1995).

Investigations and Investigating

The development of investigative skills and competencies have been identified as important outcomes for students' learning in the United Kingdom (United Kingdom Department of Education and Science, 1988) and Australia

(Australian Education Council, 1994a, 1994b). In the science profile (Australian Education Council, 1994b) *Working Scientifically* includes four phases, *Planning investigations, Conducting investigations, Processing data* and *Evaluating the findings*. Investigations are defined as activities in which students take the initiative in finding answers to problems (Jones, Simon, Fairbrother, Watson & Black, 1992). Similarly, Garnett et al. (1995, p. 27) recognise the problem solving component of investigating and define an investigation as, "a scientific problem which requires the student to plan a course of action, carry out the activity and collect the necessary data, organise and interpret the data, and reach a conclusion that is communicated in some form." In the USA scientific inquiry parallels investigating in the UK and Australia. It is described as a multifaceted activity ranging from making observations to the use of critical and logical thinking, and consideration of alternative explanations. It is claimed that "The new vision (of inquiry) includes the 'processes of science' and requires that students combine processes and scientific knowledge as they use scientific reasoning and critical thinking to develop their understanding of science" (National Academy of Sciences & National Research Council, 1996, p. 105).

Duggan and Gott (1995) consider the broad aims of science curricula in terms of conceptual and procedural understandings, a view which has similarities with science as a body of knowledge (content) and science as a method of inquiry (process) perspective (Cheung, 1994). Duggan and Gott define procedural knowledge in similar terms to investigating; "the ability of pupils to put together a solution to a practical problem from their own resources of skills and concepts rather than following a worksheet" (p. 139). They propose that procedural understandings comprise "skills" and "concepts of evidence" where skills include enabling activities such as using measuring devices and the construction of graphs and tables. These skills are consistent with aspects of *Conducting investigations* and *Processing data* described in the science profile (Australian Education Council, 1994b). On the other

hand, they define concepts of evidence to include all the concepts that are associated with obtaining evidence: variables and the idea of a fair test; measurement issues such as when to measure, what to measure and how to measure (Murphy, 1988) together with issues of data range, interval, scale and accuracy; data representation and interpretation; and validity and reliability. These aspects relate to *Planning investigations* and *Evaluating the findings* as described in the science profile (Australian Education Council, 1994b). Duggan and Gott (1995) contend that practical science should include progressively more demanding elements of procedural and conceptual understanding.

Coles and Gott (1993) have documented specific understandings and skills that may contribute to a student's ability to perform investigations and a listing of examples of these skills follows;

- identifying variables as independent and dependent,
- controlling variables and the notion of a fair test,
- distinguishing between categoric, discrete and continuous variables and the implications for different graph types,
- using equipment,
- understanding concepts of measurement including the accuracy of measurements in relation to the choice of instrument, and repeatability,
- understanding concepts of sampling including sample size and variations within a sample, the range of data, the data interval (scale), and probability,
- using tables,
- drawing the correct type of graph,
- recognising and interpreting patterns from reliable data,
- linking variables to what they represent,
- interpreting multivariate data, and
- understanding validity and reliability.

Student conceptions have been reported for some of these skills and understandings. For example, problems that students in the 12 to 14 years age group experienced in defining and operationalising a continuous independent variable were reported by Duggan, Johnson and Gott (1996). They stated that almost 50% of students defined the independent variable 'categorically' when this was inappropriate. They postulated that this shortcoming may be due to (a) a simple failure of skill at the level of using the relevant instrumentation, but they added that this was unlikely because students had the appropriate measurement skills; (b) failure to relate the continuous nature of the variable to the reality of the physical phenomenon it represents (eg. a force is represented by a push or pull); and (c) a failure to keep the whole task in mind and, as a consequence, students were unable to value the quantitative data as part of the evidence they were seeking.

The notion that students may not believe that there is a need for quantification is supported by Black (1990). He states that students only use measurement in their own problems "if they have the idea that quantification is a powerful tool and that the scientific method is powerful because it transforms problems into quantifiable form" (p. 21). Student difficulties in this area may be due to differences between the goals of science and everyday life. Reif and Larkin (1991) contend that with everyday life quantification is not important because to be 'close' is often good enough. In contrast, with science optimal strict and explicit quality control is required to measure and record data.

Another area of student difficulty is graphing. Roth and McGinn (1997) discuss three areas of science education research associated with graphing; students' interpretations of the concepts expressed in graphs, their interpretations of graphs as pictures of situations, and problems they have with the scaling of graph axes. Beichner (1990, p. 804) claims students "faired poorly when asked to explain the

concepts conveyed by the graphs," and Wavering (1989) reported that students performed poorly when asked to scale the axes of graphs.

According to Rowell and Dawson (1984) and Ross and Robinson (1987) effective schemes have been developed for teaching process skills and these can contribute to students' performances in investigations. Taken in isolation it may be argued that the skills have little if any value and that they are taught for their value in more encompassing situations (Hodson, 1992). Process skills can be taught in a holistic way as evident from a study by Roth and Roychoudhury (1993). They conducted research involving 137 boys from Grades 8, 11 and 12 physics classes and found that students developed higher order process skills through open inquiry laboratory sessions.

Specific skills and understandings are elements of a whole investigation and taken in isolation represent an atomistic view of the process of investigating. The notion of holistic teaching of investigation skills is consistent with the cognitive apprenticeship model of instruction and particularly the concept of situated cognition where learning is situated in the context of activity. In terms of learning to investigate there are calls (Duggan et al. 1996; Hodson, 1992; Roth & Roychoudhury, 1993; Woolnough, 1989, 1991; Woolnough & Toh, 1990) to adopt a more holistic approach. Some researchers (Bryce, McCall, MacGregor, Robertson & Weston, 1991; Germann & Aram, 1996; Lawson, 1995) have documented more encompassing skills that relate one phase of an investigation to another. Bryce et al. (p. 7) identify the skill to "use generalisations and observations to draw valid conclusions from the results," and Lawson (p. 53) identifies "planning and conducting controlled experiments to test hypotheses." Also Germann and Aram consider that recording data, analysing data, drawing conclusions, and providing evidence are a set of domain specific productions in which the final outcome of drawing conclusions is hierarchically related via these skills to the starting point, the

hypothesis. Because of the need to link skills and information in order to successfully perform investigations, it is arguable that an understanding of the whole investigative process is more than the sum of the parts (Hodson, 1992; Woolnough, 1991).

Some researchers (Foulds, Gott & Feasy, 1992; Germann & Aram, 1996; Kuhn, Amsel & O'Loughlin, 1988; Lawson, 1995) have commented specifically on students' abilities to connect related pieces of information. In a study involving 364 seventh grade students, Germann and Aram reported that 69% of students did not attend to the hypothesis in drawing conclusions; and that 81% did not provide specific evidence for their conclusions. Foulds et al. indicated that many students fail to see the importance of evidence in drawing conclusions. More fundamentally, students' abilities to understand the significance of their data were limited and frequently they based their conclusions on what they thought would happen regardless of their data and sometimes in direct conflict with their data. According to Foulds et al. this may point to a significant metacognitive component of science that relates to perceptions of scientific evidence in general. Foulds et al. pursue this issue under the guise of public and personal understandings. They state that evidence has a function only in the public arena because it is used to influence and shape the beliefs of other people. This perspective may have parallels with previously mentioned findings about students not perceiving the need to quantify data (Black, 1990; Reif & Larkin, 1991).

The ability to go beyond descriptions of the phenomena or data is attributed by some researchers to students' cognitive abilities. For example, Lawson (1995) claims that going beyond the phenomenological description to the testing of an hypothetical explanation is hypothetico-deductive reasoning and that this has parallels with Piaget's formal operational reasoning (Piaget, 1972). Kuhn et al. (1988) found that differentiating between, and coordinating theory and hypothesis is

a difficult task for children and some adults. They say that explicitly engaging in these science processes promotes an awareness of the need to practise the metacognitive skills required to successfully perform a scientific inquiry. They add that laboratory exercises miss this opportunity when they do not ask students to use their results to make new predictions or formulate new hypotheses.

Kuhn et al. (1988) found for young students up to Grade 9, age affects their ability to differentiate and coordinate theory and evidence. Subsequently Kuhn (1992) reported that people who had not gone to college were less able than people who had participated in a college education, to distinguish between their own explanations (theories about a phenomenon) and the evidence they had gathered. College participants were more able to reflect on their own thinking (metacognition), to view theory and evidence separately, and to evaluate the validity of theories in the light of evidence. She adds that one's prior knowledge plays a major role in generating hypotheses, designing experiments, determining what is supporting or non-supporting evidence, as well as evaluating hypotheses and revising hypotheses. Hence, "students are likely to interpret data in ways that are consistent with their personal understandings" (Germann & Aram, 1996, p. 796).

The certainty and persistence of students' beliefs are likely to be contributing influences to the ways students observe, process and interpret data and how they relate data to theory. The certainty of peoples' convictions is illustrated in a study by Kuhn (1992). With a sample of 160 people ranging in age from early adolescence to older adults, she found that "people tend to hold their theories with certainty; from one-half to three-quarters (across a range of topics) claimed that they were sure or very sure that their theories were correct" (p. 159). These findings are consistent with research on students' conceptions which indicate that students' beliefs are persistent and resilient to change (Novak, 1988; Nussbaum & Novick, 1982). Therefore it comes as no surprise that students frequently interpret their results on

the basis of what they think will happen regardless of the data before them and sometimes in direct contradiction to the data (Duggan & Gott, 1995).

Other factors also influence the way students attend to data. Rigano and Ritchie (1995) report that fudging results is common practice. They say that students make up results to get the right answer and to maintain academic results. Students fudge by matching their results with text book answers; checking their results with their peers and using the results of the peers; excluding anomalous results; and making up results. The reasons given by students for this practice are that they did not have enough time to complete the work, that they had prior knowledge of the expected outcome and matched their results accordingly, and that the equipment was poor and/or their experimental techniques were poor so their own data were inaccurate. Fairbrother and Hackling (1997) also report pressures on students to get the right answer and that there is a need to teach students more about uncertainty of measurement so that they can understand why data often differs from their expected results.

"There must be some science base to an investigation, without scientific ideas an investigation might not be scientific," (Coles & Gott, 1993, p. 8). This recommendation that investigations should have a scientific basis implies the need to link procedural understandings (processes) with concepts. Other researchers (Black, 1990; Hodson, 1992) also call for this linkage and Black (1993, p. 70) states "concept learning is inevitably involved in any science investigation". Coles and Gott contend that the "methods of investigation" are a "powerful way" to develop science knowledge and understandings (p. 8-9). They do not, however, clarify whether they are referring to developing students' (or novices') understandings of science; scientists' (expert) understandings of science; or, to developing the frontiers scientific knowledge and understandings. They also acknowledge that for students, the linking of procedural understandings with conceptual understandings is

important but add that contexts and concepts can affect the difficulty of investigations. As a consequence, they state (p. 8) that it is important that the conceptual understandings be at an "appropriate level" and caution that conceptual difficulty may act as a barrier to learning and therefore block attainment.

The notion of barriers to the attainment of investigation competencies is approached from a different perspective by others (Johnstone, 1981; Johnstone & Letton, 1990, 1991; Johnstone & Wham, 1982). Studies by Johnstone and his colleagues reveal that a student's working memory is overloaded if they have to deal with too much information in a laboratory class. This information overload results in poor learning that is characterised by a recipe following approach; concentrating on one part of the experiment and excluding the rest; busy random activity; copying the actions of others; and taking on the role of the recorder (Johnstone & Wham, 1982). Therefore, whilst there is a need to link students' understandings of the procedures and methods of investigating to conceptual understandings, caution needs to be exercised to ensure that students do not experience cognitive overload. In terms of practical work contributing to students' learning of concepts, Hodson (1993, p. 94) cautiously says the "empirical evidence concerning the efficacy of practical work as a way of learning scientific knowledge is difficult to interpret and somewhat inconclusive." However he concludes that "on balance it cannot be argued that practical work is superior to other methods and on occasions it seems to be somewhat less successful."

In terms of learning to investigate, practice and explicit instruction involving a Karplus learning cycle (Karplus, 1977), have been found to be significant factors in improving students' performances in planning and conducting investigations over an eight week period (Toh, 1991). Germann and Aram (1996) acknowledge the contribution of practice by stating, "Students need ample opportunities to practise the science processes within a variety of investigations. Together with appropriate

feedback and modelling, students should become more proficient and independent" (p. 775).

Gott and Duggan (1995) discuss factors affecting students' success at investigations. These include the science conceptual demand or difficulty; the procedural complexity which refers to the number and type of variables; the interaction of the concepts and procedural complexity; the age of the students; whether the context is scientific or an everyday context; and student factors including motivation, expectations, perception, gender and culture. They reported the results of a limited study which found that "the effect of openness was not significantly different in terms of the task score from that for the more directed, closed tasks" (p. 60).

Within this review of literature factors relating to a student's success at performing investigations have been discussed. In summary, these have included students' prior investigation skills and competencies (procedural knowledge); their knowledge of the relevant science concepts (conceptual understandings) (Coles & Gott, 1993; Germann & Aram, 1996); their abilities to associate different aspects of the investigation (Hodson, 1992); and "powerful affective attributes, including commitment and determination" (Hodson, 1992, p. 133). Factors relating to the nature of the investigation also have a bearing on students' success including, for example, the clarity of the task (Germann & Aram, 1996), the procedural complexity (Coles & Gott, 1993) and difficulty of specific skills (Duggan, Johnson & Gott, 1996); contextual factors (Coles & Gott, 1993); and the potential for cognitive overload (Johnstone, 1980; Johnstone & Letton, 1990, 1991; Johnstone & Wham, 1982).

Group Work: A Characteristic of the Laboratory Milieu

Small group work is a feature of science laboratory lessons, mainly because of the need to share equipment (Christensen & McRobbie, 1994). In this classroom setting, articulating is predominantly between students and is less frequently the expert-novice, teacher-student articulation espoused by the cognitive apprenticeship instructional model. According to researchers (Azmitta & Perlmutter, 1989; Brown & Palinscar, 1989; Doise & Mugny, 1984) group work structures have the advantage of providing students with opportunities to articulate their ideas and this social interaction benefits learning. It is claimed that group learning fosters cognitive skills, promotes social skills and imparts workplace skills (Linn & Burbules, 1993). In a discussion of the learning that occurs when students work in groups, Linn and Burbules document three distinct types of group learning; cooperative, collaborative and tutored learning. They add that these types of group learning vary in their potential to affect the intended learning outcome. They caution that with group learning, the instruction should be designed so that groups are able to access feedback, engage in debate, and have their progress monitored at regular intervals. Linn and Burbules also argue that group learning is not well suited to all tasks and educational goals. For example, they recommend that planning is best conducted by individuals; and carrying out plans is best achieved by a divide and conquer strategy in cooperative group work settings. Some of their major findings are that not all students benefit equally from group learning; that learning outcomes may not be achieved if the group is unable to access appropriate information; and that group learning may be unproductive for "learners who have dysfunctional views of group interaction." (p. 114). They add that negotiated understandings and hence collaborative learning, feature in the development of scientific understandings because ideas are postulated, criticised, and modified, and that this process leads to better understandings. Hence "students who come to participate in and appreciate such processes, therefore, benefit not only by having more sound beliefs but by

having a more accurate understanding of how scientific knowledge actually emerges" (p. 106). In addition, Siegel (1988) contends that by participating in and appreciating such processes students develop a more discursive mode of interaction that is more reasonable and open-minded.

In science, numerous researchers have analysed group work in which students performed a variety of laboratory tasks (Christensen & McRobbie, 1994; Richmond & Striley, 1996; Roychoudhury & Roth, 1996). These studies analysed group work in different ways. Christensen and McRobbie examined the social skills, discourse moves, and cognitive skills of a group of students. Richmond and Striley considered the roles that group members adopted and made links between these roles and the ease with which students developed understandings. Roychoudhury and Roth classified patterns of interaction in terms of the dominance of group members. These observational studies of group work make the point that the social and cognitive dimensions of group behaviour are "intertwined" in the sense that the development of students' understanding depends on "the interaction between the two" (Richmond & Striley, 1996, p. 843).

In an Australian study by Christensen and McRobbie (1994), four first year primary teacher education students participated in practical activities to illustrate basic science concepts and the group discourse was analysed using categories described by Barnes and Todd (1977); social skills, discourse moves, and cognitive strategies. Christensen and McRobbie (p. 56) stated that "students immediately began to carry out the physical operations without any discussion about the purposes of the activity or what it meant conceptually." They add that this did not imply that students understood the task because later the students asked each other for clarification, and that for two of the activities they did not understand the significance of the task until they had finished. They also reported that most of the questions students raised were about procedures to complete the tasks; that they

generated hypotheses only on two occasions when they were specifically called for; that they only once used observations as evidence and that most of their evidence was generated from their prior knowledge. The tone of the article by Christensen and McRobbie is one of disappointment at the lack of learning that occurred, however, the researchers conclude "students in this study believed quite strongly that practical work was vital to their understanding of concepts" (p. 58).

Richmond and Striley (1996) observed six groups of four students in a tenth grade class participating in a student designed experiment involving planning, execution and interpretation. They reported that as the students became better at developing and articulating scientific arguments their level of engagement increased, thus illustrating the entwinement of the social and cognitive dimensions. The interactions within groups were connected "in complex ways with the norms guiding group behaviour" (p. 849) and different roles emerged. They claim that "the specific roles adopted by group members were critical determinants of the ease with which students developed such understandings" (p. 843). The roles were as follows. Leaders were able to generate the group's action plan and they were classified to be inclusive, passive or alienating in style. An in-depth analysis of leaders indicated that their leadership styles were not fixed and could change. Helpers were competent individuals acting in a cooperative fashion. Active non-contributors were concerned primarily with getting by, engaging in more off-task behaviour than the leader or helper and at times challenging or ridiculing the contributions of other group members. Passive non-contributors differed from active non-contributors in that they rarely participated in group activities and often copied work from others.

Over the duration of four lessons Richmond and Striley (1996) reported that although students improved at identifying the relevant problem, collecting data, stating an hypothesis, processing the data and interpreting its meaning, areas of

difficulty were identified. The difficult areas were (a) differentiating between a problem and hypothesis, (b) understanding the value of controls in designing experiments (c) distinguishing between what they observed or measured (results) and what the observations meant (conclusions). They said most students were principally concerned with carrying out tasks and had little concern for the conceptual basis of the problem. They added that approximately 5 of the 24 students were frustrated with the other students who were task oriented, and vice versa.

Roychoudhury and Roth (1996) analysed verbal interaction patterns over a seven week period during open-inquiry physics lessons with junior high school students and found it impossible to establish a set of independent, non-overlapping categories to describe their verbal interactions. They described the "prominent features" (p. 428) of the interactions to be symmetric, in which no student monopolised the interactions and turns shifted quickly amongst group members; asymmetric in which the students' interactions were not equal; and shifting asymmetric in which dominant interactions prevailed but were not fixed in that the dominance shifted amongst students. They also stated that within the symmetric interactions students shared or negotiated understandings through either a collaborative mode or an adversarial mode. When students were not able to negotiate or reach agreed understandings the sanction of ideas come through "majority rule" (p. 431). They report that students' views about working in groups were very positive, regardless of the nature of their participation in the group. They said that most students "acknowledged the benefits of 'pooling' ideas and efforts. Negative aspects of collaborative work, as pointed out by some students, were minimal" (p. 440). They added that one of the students preferred to work alone because "he was not really the group working kind of person" (p. 440). In this study, the interactions and nature of the inquiry work conducted by the students were presented positively. The researchers attribute the difference between their classroom descriptions and those of Gallagher and Tobin (1987) to be due to the fact

that in their study students had the freedom to choose the research question and to conduct their own research; they owned their experiments and were in control of them.

Kempa and Ayob (1995) examined written answers to problems to gauge students' learning from other group members following group discussion of a science problem solving task. They concluded that there was a satisfactory level of achievement from group work, and that a significant amount of learning occurred from other students because students included in their written answers points of knowledge that had been contributed by other students during discussions. Also, they stated that a significant proportion of information appearing in students' answers had not been mentioned in group discussions. They concluded that task related comments and observations made by students in the course of group discussion represent a major shared (knowledge) resource from which the students can and do learn.

Assessment

A focus of this study is the impact that different assessment regimes (teacher assessed norm referenced, teacher assessed criterion referenced and student assessed criterion referenced) have on students' acquisition of investigation skills. The assessment regimes are considered to be one aspect of the teaching and learning program and as such they are situated in the context of the whole instructional program.

A variety of terms is used to describe assessment practices. Those relating to this study include summative, formative, norm referenced, criterion referenced,

standards referenced, student self-assessed, authentic and holistic. Brief descriptions of the meanings of these terms follow.

Summative assessment is considered to indicate how well something has been learned **after** teaching is completed (Biggs & Moore, 1993). In comparison, **formative assessment** "provides feedback to both the teacher and learner **during** the teaching process" (Biggs & Moore, p. 380), and is used by "both teacher and pupils, to modify their work in order to make it more effective" (Black, 1993, p. 49). In contrast, Gipps (1994) deemphasises the timing of these assessments and emphasises their purpose and effect. Consequently, her perception of formative assessment is consistent with the purpose espoused by Black and her notion of summative assessment is that it does not feed back into the teaching learning process but is used for grading purposes. In discussing the contribution of assessment to teaching and learning, Torrance (1995, p. 3) says the purpose of assessment is "to establish, not so much what students have learned, or have not learned, but what they *might* learn in the immediate future with appropriate help from a teacher or peer". He claims (Torrance, 1993) that formative assessment has the potential to integrate the processes of teaching, learning and assessment but adds that at present it is by no means clear "what it actually looks like at the classroom level - how it is accomplished in action - and what difference it makes if any to the culture of classroom life" (p. 341). He adds that given the claims for formative assessment in the promotion of learning and the complexities of the classroom context in which those claims must be realised, then descriptions are needed of "how teachers use routine observation and questioning of pupils in assessment, how pupils respond and how particular assessment 'incidents' or 'events' are actually accomplished through teacher pupil interactions." Calls for further trialing of both formative and summative assessment in practice are made by Black, and in a review of these assessment types, he states that historically little is documented on this subject even though the potential of formative assessment to improve student learning has been

well documented (Black, 1993; Brown & Ferrara, 1985; Brown, Campione, Webber & McGilly, 1992; Torrance 1993).

It appears that formative assessment can loosely be interpreted as any teacher judgement of a student's work that can be used to improve the teaching and learning process. Some writers have attempted to provide more structure to this interpretation. Radnor and Shaw (1995, p. 132) recognise two levels of formative assessment; "integrated formative assessment" and "structured deliberate formative assessment." The former involves "no specific assessment instruments to disrupt the teaching and learning flow but teachers are forming judgements about the pupil's ability through observation and response to set tasks." The latter involves "Teachers and pupils together, consciously and explicitly engaging in assessment activities with the purpose of helping the pupils to achieve the learning tasks set." They add that structured deliberate formative assessment is guidance through a systematic feedback structure. Harris and Bell (1994) use statements with similar meanings to describe what they term informal and formal assessment. Similarly, in reporting the characteristics of formative assessment, Bell and Cowie (1997) state that it involves recognising and responding to students learning and that it may be planned or unplanned.

Although no explicit linkage between formative assessment and the teacher guidance/coaching aspect of the cognitive apprenticeship model of instruction has been articulated in the literature, essentially they have the same purpose, to improve students' learning. In the context of the cognitive apprenticeship model of instruction, observing is the process by which teachers gather information or feedback from students and make decisions about how best to provide teacher guidance/coaching in the way of scaffolding, hints and reminders. With regard to formative assessment, Torrance (1993) identifies teacher observations and questioning as ways of gathering information about students' performances.

Similarly the contribution of questioning to gaining understandings of students' interpretations and conceptualisation from a formative assessment perspective has been identified by Gipps (1994). Interestingly, both Roth (1995) and Javela (1996), in descriptions of studies that employed the cognitive apprenticeship model, used questions as a source of information on which to base coaching decisions. Evaluations by teachers are viewed as crucial in the setting of tasks that are appropriate for the learner (Bennett & Desforges, 1985), however, Gipps cautions that these judgements may be "incomplete, fuzzy, qualitative and based on a limited range of potential criteria" (p. 130).

Lost opportunities for formative assessment and the provision of teacher guidance are alluded to by Tobin and Gallagher (1987), and Tobin (1990) in their observational studies of science laboratory work. Gallagher and Tobin (1987) say that most teachers seem to be preoccupied with management in laboratory activities and Tobin (1990, p. 414) states that students need to be provided with "opportunities to reflect on findings, clarify understandings and misunderstandings with peers, and consult a range of resources which include other students, the teacher, and books and materials."

Norm referenced assessment is "interpreted according to the performance of an individual in relation to others," (Biggs & Moore, 1993, p. 384) and "the simplest method of norm referencing is a ranking." (p. 385). Biggs and Moore, state that although the scores for many commercial tests are designed to fit a normal curve, this is not the only distribution that may be used. A weakness of the scoring of norm referenced assessment is that the unit of measurement bears "little logical relation to what has been taught" (p. 390) because it is determined by the performances of others. Also, test items on norm referenced tests must discriminate so that the scores can be spread along a normal distribution with the consequence that items that do not have a high discrimination index have no function and are usually excluded from

tests (Gipps, 1994). Biggs and Moore state that norm referenced assessment involves competition and that this may contribute to an "affective backwash" (p. 393), the effects of which are dependent on the achievement orientation of the student and hence can be positive or negative.

Following Glaser's (1963) seminal paper on **criterion referenced** assessment which served to separate educational measurement from classical psychometric testing, every development in educational measurement has stemmed from the criterion referenced model (Wood, 1986). Criterion-referenced assessment determines whether individuals meet particular task requirements (Biggs & Moore, 1993; Gipps, 1994). Claims about the strengths of criterion referenced assessment are numerous and varied:

- It enables teachers "to use teaching contexts that allow students to learn at their own pace" (Biggs & Moore, 1993, p. 386).
- Knowing the outcome of what is to be tested makes it possible to work towards the requisite level. (Biggs & Moore, 1993).
- It will improve learning because the learners and teacher are able "to judge a particular assessment against certain criteria" (Harris & Bell, 1994, p. 101).
- There is less need to discriminate with this type of assessment as there is with norm referenced assessment (Gipps, 1994).
- It allows teachers to direct their teaching to the areas on which a student does not perform well whereas with norm referenced assessment the teacher would have to look at performance item by item in order to know where to target their feedback (Popham, 1992).

Many of these views are supported by Hodson (1992) and although he welcomes the shift from norm referenced to criterion referenced assessment, he adds that this has led to demands for more precise definitions of intended learning outcomes. This is a point that is not missed by other researchers. Black (1993,

p. 58) cautions that if the "criteria are broad and vague, their formative value can be lost, whereas if they are made too specific teachers will drown in their burgeoning numbers." Gipps (1994, p. 87) describes this dilemma as a Catch 22 situation as too "loosely defined criteria limit the validity and dependability of the assessment" and "too narrowly and tightly defined assessment criteria lead to fragmentation and overburdening of discrete assessment tasks." In addition to the specificity with which the criteria are defined, other problems have been identified with criterion referenced assessment and these include:

- issues associated with reliability including a "lack of consensus about a range of adequate approaches to evaluating reliability" (Gipps, 1994, p. 84); and the
- the effect of context on performance. According to Popham (1984, p. 39) if the domain is specified then it is possible to achieve "functionally homogeneous, that is, essentially interchangeable" assessments. This idea is challenged by Gipps (1994, p. 88) as she states "research shows that context factors will critically affect pupil performance." She continues, "it is not possible therefore to decontextualise test items, nor to produce specifications so precise that it is possible to produce 'identical' test items."

Gipps (1994) sums up criterion referenced assessment as follows, "strict criterion referenced assessment is clearly unmanageable, and undesirable, particularly in an educational assessment framework" (p. 96).

Gipps (1994) refers to the move away from norm referenced assessment without going as far as criterion referenced assessment, as **standards referenced** assessment. This move is occurring in some Australian states and in the Western Australian government schools system (Education Department of Western Australia, 1997). Targets or outcomes are described in a curriculum framework and teachers are given explicit standards for judging student performance based on a developmental continuum of levels of performance. Sadler (1987, p. 193) an exponent of standards referenced assessment, claims that the quality of student work

can best be judged only by direct qualitative human judgment. He argues that standards referenced assessment is a feasible and credible form of assessment because teachers' qualitative judgements can be made dependable, provided that standards are developed and promulgated in appropriate forms and that teachers are given the relevant conceptual tools and practical training. Although "not yet proven" Gipps (p. 95) contends that this assessment model is neither norm referenced nor criterion referenced.

Broadfoot, James, McMeeking, Nuttall and Stierer (1990, p. 95) define **self-assessment** to refer to "specific judgements or ratings made by pupils about their achievement, often in relation to teacher designed categories." They state that self-assessment has three main uses. It is used to provide feedback, generally about aspects of the instructional program. The use of the term self-assessment in this context, however, is not necessarily consistent with the Broadfoot et al. definition in which self-assessment is associated with student achievement. It involves students making specific judgements to rate or estimate their achievement in relation to previously determined criteria. In addition, it is used by students in the appraisal of their performances for diagnostic purposes.

Broadfoot et al. (1990) report that evidence they have collected highlights three problems emerging from students' self-assessments; superficiality, modesty, and the persistence of norm referencing. In addressing superficiality they say "we have evidence that students are strongly influenced when making assessment judgements of themselves by their perceptions of the kinds of assessments teachers will find acceptable" (p. 96). Presumably this means that students superficially model expected teacher assessment patterns when assessing their own work rather than objectively examine the detail in their own answers. Under the heading "modesty" they reveal that students "underestimate their achievements in order not to lose face if they are subsequently down graded." They add that students' self-

assessments tend towards norm referencing because students do not have a frame of reference on which to base their performance and therefore make comparisons with others in the class.

Klenowski (1995) identified two broad types of student self-assessment processes; informal and formal. In a study conducted in Australia and the UK, she claimed that informal processes were integrated into the teaching and learning practice quickly, verbally, and pragmatically. Informal self-assessment may be associated with the self-reflective and metacognitive practices described in the cognitive apprenticeship model of instruction although this link has not previously been articulated in the literature. According to Klenowski the formal processes were more paper based with a tangible outcome because they were used to evaluate the student's progress. She contends that elements of both types of self-assessment included the use of criteria, teacher-student interactive dialogue (presumably the provision of teacher guidance through articulating), and the assignment of grades. In her study, intended learning outcomes for students included independence in their learning; responsibility for decision making relating to assignments; proactivity; and creativity in taking charge of their own work. The findings indicated that with student's self-evaluation it was possible to see an empowering impact on students.

In recent years there have been calls to assess students' science investigation competency through teacher observations made while students are performing investigations. The notions of authentic and holistic assessment underpin these calls. **Authentic assessment** is described by Biggs and Moore (1993, p. 383) as involving tasks that are devised in "situated contexts" that are designed to replicate "everyday use of the knowledge in question." **Holistic assessment** places an emphasis on the assessment of the whole task (Black, 1993). According to Hodson (1992, p. 130), holistic learning and assessment are particularly important with "doing science" because, as previously mentioned the overall task is "more than the

sum of its parts". It is likely that calls for authentic and holistic assessment are the result of similar calls for authentic and holistic learning experiences. In contrast to the advocates of the holistic assessment of laboratory tasks, Cheung (1994) claims that whether investigative tasks should be assessed holistically or in their parts remains a controversial issue which needs further critical analysis.

Numerous researchers have reported that inferences about practical ability cannot be drawn with confidence from results of written tests (Al Busaidi, Allsop & Lock, 1992; Black, 1990; Doran & Tamir, 1992; Hargraves & Lynch, 1987; Tamir & Doran, 1992). According to Black (1993, p. 63) the question about whether practical tests per se, are valid measures of practical ability needs to be asked because a "good correlation" between performances on the various component skills that make up the practical work has not been found. Gipps (1994, p. 105) says "generalisability is a particular problem for performance assessment" because "performance is heavily task dependent and we cannot take performance on one task to imply that the student could do the task in another domain." Shavelson and Baxter (1992, p. 23) claim that if performance assessment is based on tasks, "to get an accurate picture of individual science achievement the student must perform a substantial number of investigations - perhaps between 10 and 20." They say that the reason for the large number of tasks is because tasks/performances are grounded on a specific domain of knowledge, hence they are more sensitive to performance differences than more general process assessment. Similarly, Lock, (1989, 1990) showed from a factor analysis of practical tests that a coherent construct could emerge but that it was necessary to average over several tasks because of the context dependence of students' performances.

Some laboratory assessment schemes are based on checklists of criteria and teachers' observations of students' performances, and examinations of written components of students' investigations to determine the extent to which specific

criteria are met. For example, TAPS 3 (Bryce et al., 1991) identifies 14 assessment criteria and teachers assess whether students pass or fail on each criterion. From a different perspective, Mines (1995) advocates that teachers use checklists to determine whether students have the necessary skills **prior** to an investigation if the investigation is to be used for assessment. He recommends that teachers check that students have the prerequisite knowledge and practical skills; and that they know what to change and what to measure. Mines adds that "the art of skilful questioning appears to be crucial to achieve the balance between giving students suitable guidance and leaving sufficient scope for them to think independently" (p. 14).

Toh and Woolnough (1990) state that in the UK there has been wide concern about difficulties associated with practical course work assessment and, in particular, that of whole investigations. Of prime concern is the difficulty in obtaining a valid measure of the outcomes of investigative tasks without resorting to teacher observations of individual students. Woolnough and Toh (1990, p. 128-129) describe three alternative approaches to the assessment of investigations; "very tight, very reliable, very prescribed test items which are teacher-proof" or a "series of more general criteria which describe what it is to be good at doing investigative science and trust the teachers in their professional developments" or "written reports". They developed three types of report sheets on which students wrote what they had accomplished during an investigation. The written report sheets were described as uncued, broadly cued and fully cued. Students' performances were assessed by an observer and by an analysis of the different written report sheets that they completed. The findings were that there was a good correlation between the observed behaviour and the written account for each type of report form. The results for the broadly cued report sheet and the observed behaviours gave a Pearson product-moment correlation of 0.80 and the uncued and fully cued were 0.70 and 0.67 respectively. They conclude that using a "broadly cued report sheet we have a

viable method of assessing pupils' performance on a wide range of investigations" (p. 130).

In Western Australia, the location of this research, there have been moves to incorporate investigations into the science curriculum (Curriculum Council, 1997) and to assess them using standards referenced assessment called Student Outcome Statements (Education Department of Western Australia, 1997). This assessment considers eight levels of achievement for four phases of investigating; planning investigations, conducting investigations, processing data and evaluating the investigation. A copy of the standards is located in Appendix A. The assessment of students' investigations is discussed in more detail in Chapter 3.

Summary of the Chapter

This Chapter addressed literature associated with the constructivist perspective of learning, the cognitive apprenticeship model of instruction, science investigations and group work, and assessment. Several studies that involved aspects of the cognitive apprenticeship instructional model were discussed and these included some or all of the following teaching strategies; modelling, coaching, scaffolding, fading, articulating and self-reflection. The skills and competencies developed through science investigations were discussed and difficulties that students experienced were highlighted. Research on group work in science, which is a feature of laboratory classes was also discussed. Finally, different types of assessment were presented and the assessment of science investigations was addressed.

CHAPTER 3

METHODOLOGY OF THE RESEARCH

Overview of the Chapter

This study pursues three themes; the learning of investigation competencies through performing science investigations, the cognitive apprenticeship model of instruction as applied to the learning of these competencies, and the assessment of the investigations. For each of these themes research questions have been posed. This Chapter describes the setting of the research; the instructional program including organisational matters, the implementation of the cognitive apprenticeship model of instruction and the assessment procedures; the design of the study; and in conclusion, a discussion of data gathering and analysis procedures.

Setting

The study was conducted at a private school for girls in Perth, Western Australia. It was a "naturalistic" setting because although the instructional program was predetermined, it unfolded naturally within this real world setting (Patton, 1990, p. 39). Using the terminologies associated with the traditions of apprenticeship learning, the study may be described as 'authentic' research because of the natural setting, and because the findings represent the result of 'situated cognition' on the part of the researcher who is the Head of Science at the school.

The science facilities were of a high standard although additional equipment such as water baths, ukuleles and enzymes was bought for the investigations. Class sizes were relatively small (approximately 22 to 25 students) and no serious

behavioural problems were evident. The staff were well qualified, experienced science teachers. Two science technician were employed and they managed and organised the equipment for the lessons and at times trialed investigations for the researcher. The work atmosphere was pleasant and very supportive. Some, or even all of these environmental features may be present in other schools. It is acknowledged, however, that the findings of the research will be necessarily "context sensitive" (Patton, 1990, p. 39).

Students and Classes

From the Year 9 cohort three middle ability classes were selected for the study. The students' Year 8 science results were used to identify an extension class of high ability students and a very small class of low ability students. Both of these classes studied a different program from the remaining three classes. Students whose abilities were in the middle of the cohort were placed alphabetically into the three classes. It was anticipated that these classes consisted of students with comparable science abilities. The classes are named Class TN, Class TC and Class SC in this thesis. The students attended five 50 minute lessons of science each week and studied units of Biology, Physics and Chemistry each semester. The school curriculum was determined by the teachers and a tradition of ownership of the curriculum had been established because for each unit of work the teachers developed the student workbooks. Students rotated through the units of five to six weeks duration in different orders and had different teachers for some units. Students have been assigned fictitious names for reporting the research in this thesis.

Teachers

Four teachers volunteered to participate in the research. They ranged from 30 to 45 years of age and had from 8 to 28 years of teaching experience. They were

science graduates with additional teaching qualifications. The teachers were assigned to classes on the basis of their science expertise and usually taught two different classes in the same year. Fictitious names were assigned to the teachers. Mr Brogo and Mrs Grant taught physics. Mrs Cross and Miss Mills taught biology and chemistry. The rotation of classes, activities and teachers is presented in Appendix B.

Instructional Program

The instructional program involved the teaching of investigation competencies using the cognitive apprenticeship model of instruction. Students' investigations were assessed differently in the three classes. Data about the students', teachers' and researcher's perceptions of students' science investigation competencies, the cognitive apprenticeship model of instruction and the assessment regimes were gathered. As a consequence, the research may be considered to be "holistic" (Patton, 1990, p. 39) in its approach because a broad range of instructional issues is considered from a range of perspectives.

Organisational Matters

Each class completed a set of 10 lessons involving investigations during May, August and November, and therefore the instructional program comprised 30 lessons. The investigations students performed related to the science content that was taught. The program comprised two worksheets, two teacher modelled investigations and six student directed investigations. A typical lesson sequence is located in Appendix C.

All of the investigations were of a similar type, in that students were given a problem to investigate that required them to examine the relationship between an

independent and dependent variable. Students worked in groups of three and different groups chose different independent variables to investigate. Students chose their partners and where possible the same groupings were maintained for the year. The equipment was readily available in the laboratory and students selected what was required. Three 50 minute lessons were allocated to each investigation and students were advised to spend the first lesson planning and trialing, the second lesson gathering data and the third lesson processing the data and drawing conclusions.

The investigations were conceptually linked to the intended conceptual learning outcomes of the Biology, Chemistry and Physics units that the students studied and it was intended that the investigations would reinforce and consolidate these outcomes. It was considered, also, that students and teachers would perceive the investigations to be authentic activities if linkages to the conceptual units were evident. Further, it is purported (Collins et al., 1989) that situating activities (investigations) in the context of their use helps to solve problems of brittle and inert knowledge.

The biology investigations included the lipase investigation and trypsin investigation (Appendix D); the chemistry investigations included the catalyst investigation and the Panadol investigation (Appendix E); and the physics investigations were the ukulele investigation and the electromagnet investigation (Appendix E). The lipase, trypsin, electromagnet and catalyst investigations were developed from ideas presented in *TAPS 3: How to Assess Open-ended Practical Investigations in Biology, Chemistry and Physics* (Bryce, McCall, MacGregor, Robertson & Weston, 1991). The ukulele investigation was developed by Garnett, Davies and Hackling (1996).

The lipase investigation was used as a pretest investigation and the trypsin investigation was the posttest. For both of these investigations students were required to complete an Investigation Planning and Report Sheet (IPRS) and

condensed versions of these are located in Appendix D. For the other investigations scaffolded IPRSs were used and a condensed version is in Appendix F. The notion of providing direction through scaffolding the IPRSs is consistent with the cognitive apprenticeship model of instruction in which scaffolding is advocated to provide help and support for students as they carry out tasks (Collins, Brown & Newman, 1989). The final IPRS was less scaffolded and this represented fading or the gradual withdrawal of support. The IPRS which was used for the pretest investigation necessarily matched the posttest IPRS.

Two worksheets were used. One of these, the worksheet on terminology, was designed to help students master the terminology and concepts associated with investigations such as hypothesis, independent variable, dependent variable and to identify the factors that needed to be controlled in investigations. The other worksheet on marking an investigation, was designed to help students reflect on how investigations were assessed and to become aware of typical errors that Year 9 students made when writing-up investigations. Condensed versions of the worksheets are located in Appendix G. The worksheet on terminology was developed by the researcher and M. Hackling, and the worksheet on marking an investigation was developed by the researcher and based on an investigation that the preceding Year 9 students had completed. Two teacher modelled investigations were developed by the researcher, the acid and calcium carbonate investigation and the pitch of a closed pipe investigation (Appendix H), and were implemented before the catalyst and ukulele student investigations respectively.

The order in which classes completed the activities is presented in Appendix B. During May, Class TN and Class TC completed the Test of Science Investigation Skills (TOSIS), the worksheet on terminology, the pretest (lipase) investigation, the teacher modelled acid and carbonate investigation and the catalyst investigation. During August the activities were the teacher modelled pitch of a

closed pipe investigation, the ukulele investigation, the worksheet on marking an investigation and the electromagnet investigation. During November these classes participated in the Panadol investigation and the posttest (trypsin) investigation. The order for Class SC was slightly different. They participated in the TOSIS, the worksheet on terminology, the pretest (lipase) investigation, the teacher modelled pitch of a closed pipe investigation and the ukulele investigation. During August Class SC participated in the teacher modelled acid and carbonate investigation, the catalyst investigation the worksheet on marking an investigation and the Panadol investigation. Finally, during November they participated in the electromagnet investigation and the posttest (trypsin) investigation.

Implementing the Cognitive Apprenticeship Model of Instruction

The cognitive apprenticeship model of instruction was used as the instructional model to facilitate students' development of investigation competencies. Features of the instructional model have been discussed in detail in the Literature Review (pp. 10-19) and the interpretation of the components of the model used in this study are presented in Chapter 1 (p. 4). Briefly, the instructional model includes modelling investigations to make tacit knowledge explicit, providing students with help by coaching and scaffolding in the zone of proximal development (Vygotsky, 1986), gradually reducing or fading the level of support offered to students, encouraging students to articulate as they learn, and encouraging students to develop the skills of self-reflection and metacognition. Teachers received two hours training on the instructional model during which time these features were discussed. In addition, the teachers spent several hours trialing the investigations to identify problems that may arise.

The teacher modelling aspect of the cognitive apprenticeship model of instruction was achieved when students observed their teachers model two

investigations. This interpretation of modelling is consistent with global modelling (Javela, 1996). It was intended that the modelled investigations would address the following investigation competencies; planning investigations, conducting investigations, processing data and evaluating investigations. The modelled investigations were situated in diverse settings (physics and chemistry) with the intent that students would observe how investigation skills and competencies were applied in varied contexts. It was anticipated that the emphasis on specific competencies would vary with the domain of the investigation, and that teachers would expose and share their knowledge, skills and competencies in different domains. Coaching was achieved by the teachers helping students as they participated in the investigations, although it was not undertaken during the pretest and posttest investigations. Scaffolding was provided by scaffolded IPRSs (Appendix F) that students were required to complete as they performed all investigations except for the pretest and the posttest investigations. Fading was achieved by withdrawing the modelled investigations and by using the IPRS with less scaffolding for the posttest investigation. Articulating was achieved through teacher-student interactions and student-student interactions. Self-reflection and metacognition were encouraged when students completed the last page of the IPRS, the worksheet on marking an investigation and the student questionnaires.

Assessment

Students received two levels of formative assessment or coaching during the instructional program; informal and formal (Harris & Bell, 1994). The informal formative assessment comprised teachers questioning and helping students as they performed the investigations and was based on teachers' observations of students as they worked, and questions that students asked the teachers. The formal formative assessment was deliberately structured to occur after students had completed an investigation. These assessments have the potential to improve students'

performance on subsequent investigations (Black, 1993; Gipps, 1994). The formal formative assessment regimes for each of the classes differed and they are described as follows. The researcher explained and discussed the different assessment procedures individually with teachers before and during the program of instruction.

Teacher assessed norm-referenced assessment: Class TN

Class TN was teacher assessed and the overall grades were norm referenced. The assessment instructions for the teachers are presented in Appendix I. The students in this class were not provided with the assessment criteria, however they were provided with an assessment sheet on the final page of their investigation that was completed by their teacher. Teachers ranked the students in order of merit and assigned an A grade for three students, a B^+ grade for five students, a B grade for eight or nine students, a C^+ grade for five students and a C grade for three students. The teachers also wrote individual comments about a student's performance on her investigation planning and report sheet. Following the assessment of the students' investigations, the teacher devoted 20 minutes to whole-class feedback and then students were allocated 10 minutes to make corrections or improvements to their investigation planning and report sheets using a different coloured pen.

Teacher assessed criterion referenced assessment: Class TC

For this class the assessment was teacher assessed and criterion referenced. Students were provided with the assessment criteria on the last page of their IPRS prior to performing the investigation (Appendix I). Teachers assessed the performance of the students according to each criterion and assigned an A, B or C grade depending on whether the criterion was fully met, partially met or not met and an NR grade was assigned when a criterion was not relevant. Teachers, at their own discretion, wrote a comment next to each criterion. The instruction for feedback to students after they had received their assessed investigation was similar to Class TN.

Student assessed criterion referenced assessment: Class SC

For this class students assessed their own work according to the same criteria that the teachers of Class TC used. Students were given the assessment criteria prior to performing the investigation (Appendix I). They were instructed to compare their performance with the researcher's responses on a master answer sheet and to make judgements about the degree to which they satisfied each criterion. Numerous answer sheets were prepared to cater for the different variables that the students could choose to investigate. Students were instructed to assign grades for each criterion in a similar way to the teachers of Class TC. Students allocated 30 minutes to this procedure and during this time they were asked to also make corrections or improvements to their IPRS.

Design

A qualitative research paradigm was selected as being most appropriate for the research because it has the potential to provide "deep insights" and capture direct quotations of personal perspective and beliefs (Patton, 1990, p. 39). The study was a naturalistic inquiry which embraced notions of exploration, description, illustration, realisation (in the sense that events are made real to the audiences) and the testing of research questions (Guba & Lincoln, 1989). Clearly the research agenda was not genuinely open because answers to research questions were sought in relation to the themes of the research. An interpretative research methodology (Erickson, 1986) was used to seek patterns in the data so that the data could then be related to the research questions.

Opportunities for "personal contact and insights" were enhanced because the researcher worked at the school where the research was conducted, however, an "empathetic stance" was adopted in which a non-judgemental approach was

exercised (Patton, 1990, p. 39). This was an important and critical feature of the research. It was considered by the researcher that breeching this empathetic stance would destroy the teacher trust on which the research was grounded and would not serve to answer the questions at the heart of the research.

Care was taken to overcome some pitfalls of qualitative research (Roberts, 1996). Roberts contends that too often researchers present case stories of descriptions instead of case studies with findings. Procedures characteristic of interpretive research methodologies (Erickson, 1986) were used to overcome this weakness. Assertions were generated from various data sources and were numbered and subsequently collated to develop more general assertions which represented the findings of the research. At times, attempts to quantify the assertions were made with the use of the words *some* or *most*, with the former referring to less than 25% of the sample and the latter to more than 75%.

Roberts (1996) also states that there is often insufficient utilisation, presentation, and/or discussion of appropriate conceptual frameworks. In this study this weakness is addressed by discussing the findings in relation to the instructional model and conceptual framework (Chapters 9 and 10). In addition, Roberts claims that often qualitative studies fail to reach a conclusion and that this short changes the reader and threatens "the very fragile niche that sound qualitative research has established" (p. 248). In this study research questions are addressed (Chapter 10) and conclusions are drawn in relation to these questions.

Figure 2 provides an overview of the linkages between the data and assertions. From the various data sources (pretests and posttests, student questionnaires, teacher interviews, student group interviews, and audio and video recordings) assertions were formulated and many of these were triangulated across the data sources. As well, a search for disconfirming data was conducted to ensure

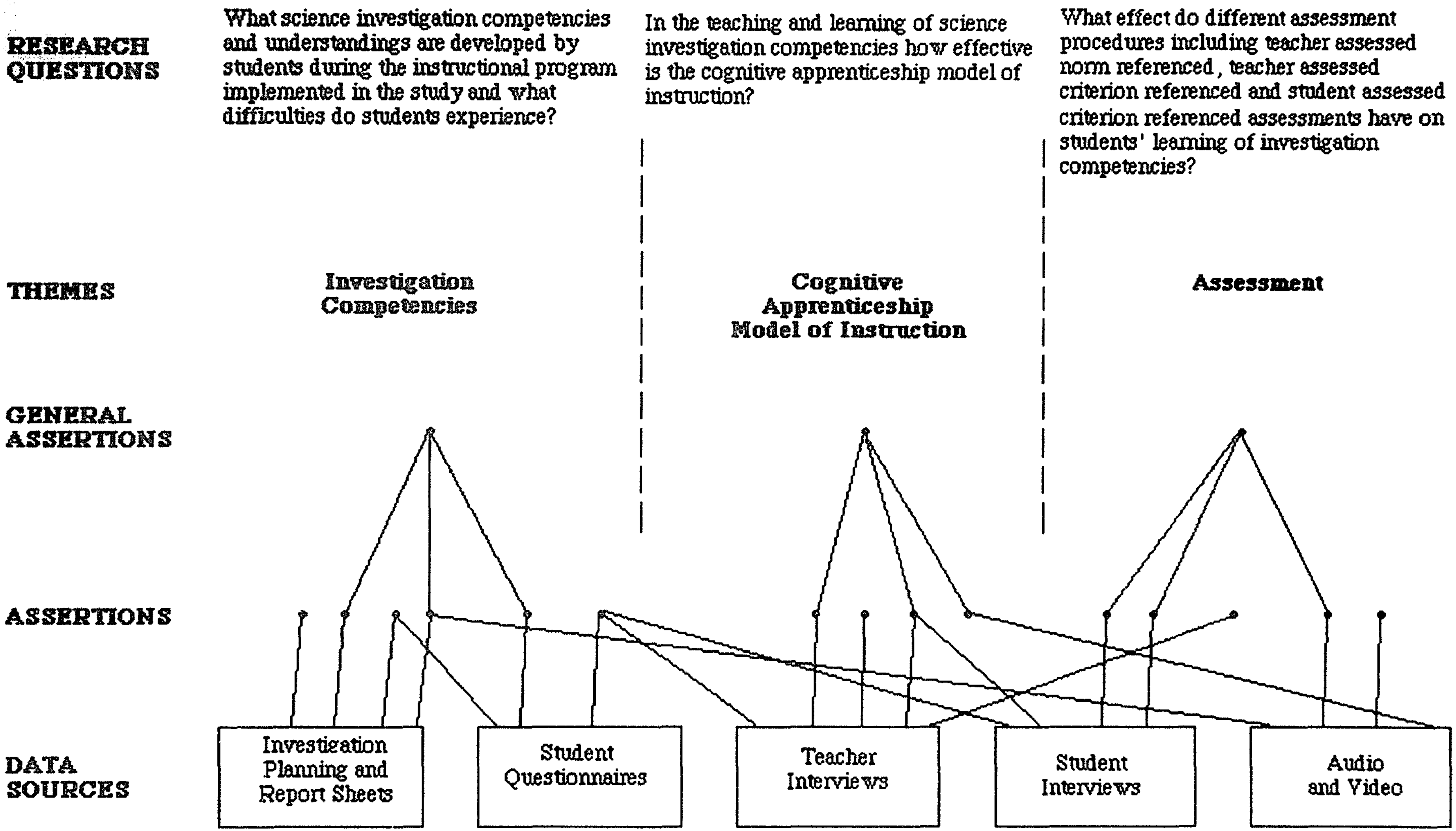


Figure 2. A schematic representation of the association between the data sources, assertions, general assertions and themes of the research

that inconsistencies in the assertions generated from different sources were identified. The assertions were clustered into the themes of the research and general assertions or findings were formulated in response to the research questions. Therefore, through a process of inductive analysis, the researcher became immersed in the details and specifics of the data to discover dimensions and interrelationships (Patton, 1990). Erickson (1986, p. 148) similarly describes the researcher looking for "key linkages" which lead to "patterns of generalization within the case at hand, rather than from one case or setting to another." He cautions that in this process "rare events are not handled well" and that "frequently occurring events can come to be better understood than can rare events" (p. 148).

Multiple data sources and the triangulations of data have been advocated as ways to secure in-depth understandings of the phenomenon in question (Brewer & Hunter, 1989). In this study viewpoints of the students, teachers and researcher were brought to bear on particular situations. This triangulation process is not viewed as a tool or a strategy of validation but as an alternative to validation (Denzin, 1989a, 1989b; Fielding & Fielding, 1986; Flick, 1992). Flick contends that "the combination of multiple methods, empirical materials, perspectives and observations in a single study is best understood, then, as a strategy that adds rigor, breadth, and depth to any investigation" (p. 194). In this research the credibility and trustworthiness of the data is improved by an audit trail of the assertions. This is addressed in the data gathering and analysis section which follows.

Some quantitative data procedures were conducted to determine the effectiveness of the instructional program and the assessment procedures. From a science perspective, Black (1990, p. 21) states that "quantification is a powerful tool" of scientific research. Within education research methodologies, there is debate about mixing qualitative and quantitative research paradigms (Cook & Reichardt, 1979; Fetterman, 1988; Filstead, 1970; Lincoln & Guba, 1985; Patton, 1986; 1988). Guba

and Lincoln (1988) have argued that the internal consistency and logic of each approach or paradigm mitigates against methodological mixing of different inquiry modes and data collection strategies. According to Patton, (1990, p. 193) "their cautions are not to be dismissed lightly. Mixing parts of different approaches is a matter of philosophical and methodological controversy." Patton (1981, 1990) concludes that the practical mandate in evaluation, to gather the most relevant possible information, outweighs concerns about methodological purity based on epistemological and philosophical arguments. Denzin (1978) and Denzin and Lincoln (1994) support the pragmatic approach adopted by Patton and identify methodological triangulation as the use of multiple methods to study a single problem.

Data Gathering and Analysis

In this research written documents included students' pretests and posttests, portfolios of worksheets and Investigation Planning and Report Sheets, and student questionnaires. Semi-structured interviews were conducted with groups of students and individually with teachers, and observations were made by the researcher from audio and video tapes. This approach is consistent with Patton's (1990) ideal-typical qualitative method strategy which comprises three parts; (1) written documents yielding excerpts, quotations and entire passages from open-ended written responses to questionnaires; (2) in-depth, open-ended interviews consisting of direct quotes of experiences, feeling, opinions and knowledge; and (3) direct observations consisting of detailed descriptions of students' and teachers' activities, behaviours and actions, and the full range of interpersonal interactions and organisational processes that are part of observable human experience.

Some of the written documentation from the pre and posttest and student questionnaires produced quantitative data. These data were considered to be

appropriate for providing insights into the effectiveness of the instructional model and the assessment procedures. The resulting methodological triangulation of qualitative and quantitative data was consistent with the pragmatic approach to research described by Denzin (1978), Denzin and Lincoln (1994), Patton (1981, 1990) and McFee (1992).

Test of Science Investigation Skills (TOSIS)

The Test of Science Investigation Skills, TOSIS, (Appendix J) was developed by Hackling and Garnett (1993) as a pencil and paper test of science investigation knowledge and skills. Two parallel versions of the test were developed, one in a biology context and the other in a physics context. In this research the biological version was used because students started and completed the instructional program with biology investigations. The test reliability is reported to be 0.77 (Hackling & Garnett, 1993).

The test comprises five open-ended questions set in a biological context and seven multiple choice questions that are context free. A description of the skill or concept area is shown in Figure 3, together with the question number/s and the subtest score for each particular skill or concept area. The performances of the classes on the TOSIS are compared on the total scores and on subtest scores corresponding to skill and concept areas (Chapter 4).

The scoring procedures for the TOSIS, including the scoring for each question, are presented in Appendix J. The scoring of the multiple choice questions 6 to 12 was objective and either right (one mark) or wrong (zero marks). The scoring of question 4 and question 5 proved to be objective as students were either right or wrong, or had omitted a response. The scoring of questions 1, 2 and 3 was more difficult and the following procedures were adopted to ensure reliable scoring.

Identifying variables *Score = 6*

Q. 1 Identifies three variables for testing from an investigation scenario.

Q. 6 Identifies the independent variable from a description of an experiment.

Q. 7 Identifies the dependent variable from a description of an experiment.

Q. 9 Identifies the controlled variables from a description of an experiment.

Writing an hypothesis *Score = 1*

Q. 2 Writes an hypothesis as a relationship between an independent and dependent variable.

Planning an investigation *Score = 8*

Q. 3 Plans the design of an investigation with consideration of control of variables, how the dependent variable will be measured, and sample size.

Drawing conclusions *Score = 2*

Q. 4 Formulates an appropriate conclusion.
Recognises that the experimental design is poor and therefore conclusions must be tentative.

Identifying methodological limitations *Score = 2*

Q. 5 Recognises that a key variable was not controlled.
Recognises that the sample size was not adequate.

Understanding the concepts of hypothesis, theory, data and conclusions *Score = 4*

Q. 9 Selects a definition of an hypothesis as "a tentative explanation that can be tested".

Q. 10 Recognises that experimental results should be "accurate and repeatable".

Q. 11 Selects a definition of a theory as "explanations that have been supported by the results of many experiments".

Q. 12 Recognises that scientific conclusions are "subject to revision".

Figure 3. Description of investigation skill and concept areas assessed by the TOSIS and arranged into subtests corresponding to skill and concept areas

For question 1 students were required to identify three variables. Most identified at least two variables that were relatively easy to control and measure, such as soil type, fertiliser type, the amount of fertiliser, the amount of water, pot size, and pots or no pots. The third variable selected by students was often one that would make it difficult or impossible to conduct a rigorously controlled experiment. Such variables included the amount of sunlight, the climate or weather, and the position or place where the experiments were conducted. These variables were accepted as correct because students had not been asked to think of ways to measure the variables. Unacceptable variables were the "conditions" and the "environment" because it was believed that these were too vague.

The accepted responses to question 2 included declarative statements of a relationship between the independent and dependent variables, such as "The more fertiliser the more cucumbers will be produced." Many students said that the cucumbers would "grow better" therefore giving no indication of how the dependent variable would be measured. These hypotheses were accepted because the criterion of the hypothesis being a testable statement was not required. For example "The cucumber grows best in clay soils," was marked as an acceptable response. Unacceptable responses included those that, (a) outlined a method such as, "Put fertiliser in one pot and none in the other and see which cucumber grows faster"; (b) posed a question such as "Do yellow cucumbers grow faster in clay or loam?" and (c) included a theory such as "Yellow cucumbers grow better outside because they are more exposed to the elements" and also, "The bigger the pot the larger the root system so the bigger the plant."

For question 3, responses were allocated four marks when students indicated that they planned to control four variables. Although it would be most unlikely that the variables sunlight/light and greenhouse or no greenhouse could result in rigorously controlled experiments, these responses were accepted because students

were not asked how to control the variables. Three marks were allocated when students indicated that they planned to control three variables. Students were allocated two marks if they indicated that two variables would be kept constant or if the conditions needed to be kept the same. Question 3(g) was marked quite rigorously. Students were expected to state how they would measure the yield. They needed to refer to the number of cucumbers and therefore a tally or count was accepted. Responses that implied that students would record the results and/or observe what had happened were not accepted.

The TOSIS was administered as a pretest and posttest to determine whether students had improved their understanding of the terminology and concepts associated with investigations. A two-way ANOVA was conducted to determine if there were an effect for test occasion, group and/or interaction between the two. The results of the TOSIS are presented in Chapter 4, Tables 1 and 2.

Problems associated with the validity and authenticity of pencil and paper tests as measures of laboratory performance have been raised in the Literature Review (p. 42) and for this reason students also performed pretest and posttest investigations to determine the effect of the instructional program.

Investigation Planning and Report Sheet (IPRS)

An Investigation Planning and Report Sheet was designed to assess students' science investigation skills in pretest and posttest investigations (Appendix D). The IPRS was broadly/partially cued and similar planning and report sheets have been found to have a Pearson product-moment correlation of 0.80 with students' actual performances recorded on a check list by an observer (Toh & Woolnough, 1990). Also, in a resource book prepared for the Education Department of Western Australia, Hackling (1998) indicated the value of using partially cued worksheets to

elicit responses from students regarding planning and conducting investigations, analysing data, and evaluating investigations. Teachers can make judgements about a student's performance from the IPRS and assign levels of achievement described in the Student Outcome Statements (Appendix A). The data obtained from the IPRS are authentic in that they describe a student's thinking and doing as they perform an investigation in a group of three students. The results are likely to be affected by the performances of other group members because collaboration occurs during the planning and conducting phases of the investigation. In comparison, the TOSIS may be more reliable in assessing individual student's pencil and paper skills and understandings of investigations, however, it lacks authenticity because it is unrelated to performance on an actual investigations. Together these data gathering procedures are complementary in nature.

The format of the IPRS for the pretest and posttest investigations was the same (Appendix D). The problems that students were asked to investigate were different, but similar in many respects. The effects of task and domain specificity on performance/practical assessment have previously been documented (Haertel 1993; Linn, Baker & Dunbar, 1991; Shavelson & Baxter, 1992) and so efforts were made to construct investigations that were parallel in content and structure. The pretest and posttest investigation problems are presented in Figure 4 and a comparison of the features of the investigations is presented in Figure 5. The features include the science content knowledge or the conceptual context, the independent variables that the students could choose to manipulate, the dependent variable and how it was measured, and the type of graph that would be needed to interpret patterns in their data.

Pretest Investigation	**Lipase** is an enzyme found in the small intestine. Lipase helps to digest food by breaking fats into fatty **acids**. For example it breaks down the fat in milk into fatty **acids**. The environment in the small intestine where the lipase works, is very slightly **basic** and the temperature is 37 °C. The fatty acids that are produced by the action of the lipase make the solution acidic. (a) What variables might affect how quickly this will take place? (b) Which variable will you investigate? (c) Write an hypothesis for your investigation.
Posttest Investigation	**Trypsin** is an enzyme found in the small intestine. Trypsin helps to digest food by breaking down protein. The environment in the small intestine, where the trypsin works, is very slightly **basic** and the temperature is 37 °C. Trypsin for example, breaks down the protein called gelatin. Gelatin is found on undeveloped photographic film and when a film is put in a trypsin solution the gelatin reacts. (a) What variables might affect how quickly this will take place? (b) Which variable will you investigate? (c) Write an hypothesis for your investigation.

Figure 4. Pretest and posttest investigations (lipase and trypsin)

It is also acknowledged that the generalisability of performance assessment is influenced by the type of instruction relating to the performance. To ensure that students were given similar directions/instruction, for both investigations, the teachers spent approximately 10 minutes of the first lesson discussing a background information sheet (Appendix D) relevant to the action of the enzyme that the students were investigating; lipase on milk for the pretest and trypsin on photographic film for the posttest. Antecedent instruction was therefore guided by the information sheet in recognition that this is an important influence on performance assessment (Gipps, 1994). After this period of instruction, students were advised to spend the remainder of the first lesson (40 minutes) planning the investigation and conducting some preliminary trials, the second lesson collecting

the data, and the third lesson finishing off their work and writing up the investigation. They worked in groups of three for the two lessons that involved planning, trialing and data collecting. During the third lesson they worked individually to write up their IPRS. Group membership was determined by the students and where possible remained the same for the duration of the program. Teachers were instructed to not help students during the pretest and posttest investigations.

Features of investigation	Pretest (lipase investigation)	Posttest (trypsin investigation)
Conceptual context	Enzyme action, provided on worksheet	Enzyme action, provided on worksheet
Independent variables that could be chosen	Temperature of milk [c] Amount of lipase [c] Amount of milk [c] Type of milk [d]	Temperature of solution [c] Amount of trypsin [c] Amount of film [c] Amount of base [c]
Dependent variable	Time for the milk to become acidic	Time for the film to become clear
Measured by	Observation of indicator colour change	Observation of film to become clear
Graph	Column or line depending on the variable selected	Line

Figure 5. Comparison of the features of the pretest and posttest investigations (Note. [c] continuous variable, [d] discrete variable)

Students' performances on the Investigation Planning and Report Sheets were coded into standards or levels of achievement using a version of the Working Scientifically strand of the Science Profile (Australian Education Council, 1994b) that was revised by the Education Department of Western Australia (1997) in a trial of Student Outcome Statements (Appendix A: Student Outcome Statements: Standards for the Assessment of the Pretest and Posttest Investigations). This form of assessment is advocated by Sadler (1987) and described by Gipps (1994) as a

move away from normative assessment without going as far as criterion referenced assessment. The performance standards represent eight levels of achievement. According to Baker et al. (1991), specifying the cognitive demands, scoring criteria, performance standards and context quality, ensures the compatibility of performance assessments.

Students were assigned levels of achievement for each of the four substrands of an investigation; *Planning investigations*, *Conducting investigations*, *Processing data* and *Evaluating investigations*. A student's level of achievement for *Planning investigations* was determined from their responses to the first three sections on the IPRS; the Problem, Planning, and Preliminary trials. Their level of achievement for *Conducting investigations* was determined from their response to the fourth section, Collecting data. *Processing data* was determined from the fifth section (Processing data) and *Evaluating investigations* from the student's responses to the last section of the IPRS (Evaluating findings). For students to be assigned a particular level of achievement they needed to exhibit competency for all aspects of the level statement that could be exhibited by performing the task. Students who did not write comments for sections of the IPRS that corresponded to particular substrands were assigned a score of zero for that substrand.

To ensure that the coding of the substrands of the IPRSs into levels of achievement was reliable, the following procedures were adopted. A sample of five IPRSs was randomly selected and independently scored by two markers. Discrepancies in the scores were discussed and a consensus was reached on individual student's levels of achievement. Following this training, a sample of 15 randomly selected investigations was scored independently by the same two markers. The proportion of judgement cases in which both markers assigned the same level for a substrand was calculated. An inter-rater reliability of 0.82 was achieved on the assignment of levels of achievement. These data are presented in

Appendix K. One of the markers then scored all of the pretest and posttest IPRSs. To determine the intra-rater reliability, one week after completing this scoring, this marker rescored 10 of the IPRSs and an intra-rater reliability of 0.88 was obtained. These data are presented in Appendix L.

The numbers of students who achieved a particular level on the pretest and posttest investigation were tallied for each of the phases of investigating (Table 3). A Wilcoxon Matched-Pairs Signed-Ranks test was conducted to determine the effect of test occasion, and following that a Kruskal-Wallis One-Way Analysis of Variance was conducted to determine if there were a statistical difference between the performance of the groups, Classes TN, TC and SC on the pretest, and on the posttest.

Students' written responses to the pretest and posttest IPRS were examined in detail for trends and patterns in the data. A number of assertions were generated from these data and they are discussed and presented in Chapter 4. At the completion of Chapter 4, the Summary of the Chapter collates the assertions.

Questionnaires

The students responded to three questionnaires at the completion of the May, August and November sequence of lessons (Appendix M). The purpose of the questionnaires was to gather information about students' perceptions of the investigations, the cognitive apprenticeship model of instruction and the assessment procedures. Some quantitative data were gathered and these are presented as descriptive statistics (Chapter 5). Students' responses to open-ended questions were collated and categorised, and some direct quotations are included. Assertions that summarise the data and are presented throughout Chapter 5. A search for data that confirmed or disconfirmed assertions grounded on previously presented data was

made. Chapter 5 concludes with a Summary of the Chapter in which a collation of assertions from the questionnaire data is presented.

Interviews

In-depth, semi-structured interviews were conducted by the researcher with teachers and groups of students participating in the study. The interviews were shaped by the tasks that the students and teachers participated in, and by the nature of the assessment. The questions were not predetermined although the researcher deliberately sought data about the instructional program and the assessment. Patton, (1990, p. 280) describes this as a "general interview guide approach" in which a set of issues that are to be explored are outlined before interviewing begins. He adds that the issues need not be taken in any order and the actual wording of the questions is not determined in advance.

Individual **teacher interviews** were conducted after each set of investigations. The purpose of the interviews was to provide the teachers' perspective of the students' investigation competencies, the cognitive apprenticeship model of instruction and the assessment procedures. Their responses were collated and presented as assertions in Chapter 6. A search was made for disconfirming and confirming data to obtain triangulation of the assertions. In Chapter 6, the Summary of the Chapter collates the assertions and documents support for previously generated assertions.

For the **students' interviews** an interactive group interviewing strategy was used (Tikunoff & Ward, 1980). Tikunoff and Ward espouse that putting participants together as a team increases both the meaningfulness and the validity of the findings. They add that the research is less intrusive and reduces the reactivity of the participants. Participants are more cooperative and have an improved understanding

of the research process. Group interviews are described by Fontana and Frey (1994) as the systematic interviewing of several individuals simultaneously in formal or informal settings. They contend that the interviewer can direct the interaction and inquiry in either a very structured or very unstructured manner.

One group of three students from each of the three classes was interviewed after each 10 lesson sequence. Three class teachers selected a particular group from their class and Groups TN, TC and SC were interviewed following each of set of investigations. All interviews were audio taped and students' responses were collated into the themes of investigation competencies, the cognitive apprenticeship model of instruction, and the assessment regimes. In Chapter 7 the groups' perceptions were summarised and reported as assertions. These data were cross referenced with previously presented assertions and consistencies and inconsistencies recorded, and collated in the Summary of the Chapter.

Observations

The purpose of observational data is described by Patton (1990, p. 26) as "to take the reader into the setting that was observed," and for this to occur he adds that the data must have depth and detail to be sufficiently descriptive for the reader to understand what occurred and how it occurred. Janesick (1994) cautions that the analysis and interpretation of observational data should be balanced by descriptions. This perspective echoes the balance previously cautioned by Roberts (1996) about qualitative research producing case stories instead of case studies.

In this research direct observations of the classes were not made because the teachers were uncomfortable with the presence of an observer. Audio and video recordings were made of the teachers and specifically Groups TN, TC and SC. Subsequently, the recordings were analysed by the researcher in terms of the

cognitive apprenticeship model of instruction and the assessment procedures. They were viewed in conjunction with students' portfolios of work. The audio and video data are presented as assertions in Chapter 8. A search was made for data that were consistent or inconsistent with previously generated assertions. The Summary of the Chapter collates the assertions that were generated and highlights confirming and disconfirming data.

Chapter 9 collates the assertions from the various data sources under the headings Investigation Competencies, the Cognitive Apprenticeship Model of Instruction and Assessment. Each assertion is first identified by the data source from where it originated and also other data sources that provide confirming evidence. Disconfirming or contrasting assertions are presented in italics. This triangulation and audit trail of the assertions was conducted to improve the credibility and trustworthiness of the data.

Summary of the Chapter

This Chapter addressed the methodology of the research. The setting, including students and classes, and teachers was described, and factors associated with the implementation of the instructional program such as organisational matters, implementing the cognitive apprenticeship model of instruction, and the assessment procedures were addressed. Methodological issues associated with the design of the research were discussed and these include data interpretation and triangulation, and the inclusion of quantitative data into a study based on a qualitative research methods. Data gathering and analysis included pre and posttest data from the pencil and paper Test of Science Investigation Skills and students' performances on investigations as measured by Investigation Planning and Report Sheets. Data were also gathered and analysed from student questionnaires, teacher interviews, student group interviews and observational data from audio and video tapes, and these procedures are discussed. Procedures implemented to triangulate the data were also addressed.

CHAPTER 4

PRETEST AND POSTTEST MEASURES OF STUDENTS' INVESTIGATION COMPETENCIES

Overview of the Chapter

This Chapter discusses pretest and posttest data that were gathered to monitor the development of students' investigation competencies. From these data inferences can be made about the effectiveness of the instructional approach and the effect of different assessment procedures implemented in the study; teacher assessed and norm referenced (Class TN), teacher assessed and criterion referenced (Class TC), and student assessed and criterion referenced assessment (Class TC).

Two complementary data gathering procedures were used, the pencil and paper Test of Science Investigation Skills (TOSIS) and practical investigation performance as recorded on an Investigation Planning and Report Sheet (IPRS). A copy of the TOSIS and the coding sheet are located in Appendix J, and condensed versions of the pretest and posttest IPRS are in Appendix D.

The Chapter first compares the pretest and posttest total TOSIS scores for the classes in the study, and the pretest and posttest scores for specific investigation skill and concept areas comprising the TOSIS. Second, it addresses students' planning, conducting, data analysis and evaluation of their investigations as recorded on the IPRS. Quantitative analyses of the classes' pretest and posttest performances are presented, and qualitative data about the nature of students' responses on the IPRS are discussed under the headings Planning investigations, Conducting investigations, Processing data and Evaluating investigations.

Test of Science Investigation Skills (TOSIS)

Features of the TOSIS and the scoring procedures, as well as information about its reliability are described in Chapter 3. Analyses of the TOSIS data were at two levels; a comparison of the pretest and posttest total scores and comparisons of the scores on specific skill and concept areas.

Total TOSIS Scores: A Comparison of Classes on Pretest and Posttest Scores

Descriptive statistics for the mean total TOSIS score for each class are presented in Table 1 and Figure 6. The data indicate that gains were achieved in the mean total test scores by each class. Using an SPSS computer program, a two-way Analysis of Variance (ANOVA) was conducted to determine whether there was an effect for test occasion (pretest and posttest), an effect for the group (Classes TN, TC and SC), and/or an interaction effect between the two. The analysis revealed an effect for test occasion, $\underline{F}(1, 2) = 66.68$, $p < .01$, no effect for the group, and no interaction effects. Therefore, it may be concluded that students' performances improved significantly over the period of intervention that employed a cognitive apprenticeship approach to instruction. In addition, there was no significant difference between the classes that were exposed to the different assessment procedures.

Table 1. Means and standard deviations for the pretest and posttest total TOSIS scores, and the gains in the means for Classes TN, TC and SC (Maximum score = 23)

Class	Pretest		Posttest		Gain in
	Mean	St dev	Mean	St dev	mean
TN (n = 23)	11.30	2.87	15.43	2.52	4.13
TC (n = 22)	11.23	3.02	15.45	2.97	4.22
SC (n = 24)	11.13	3.42	14.29	3.88	3.16

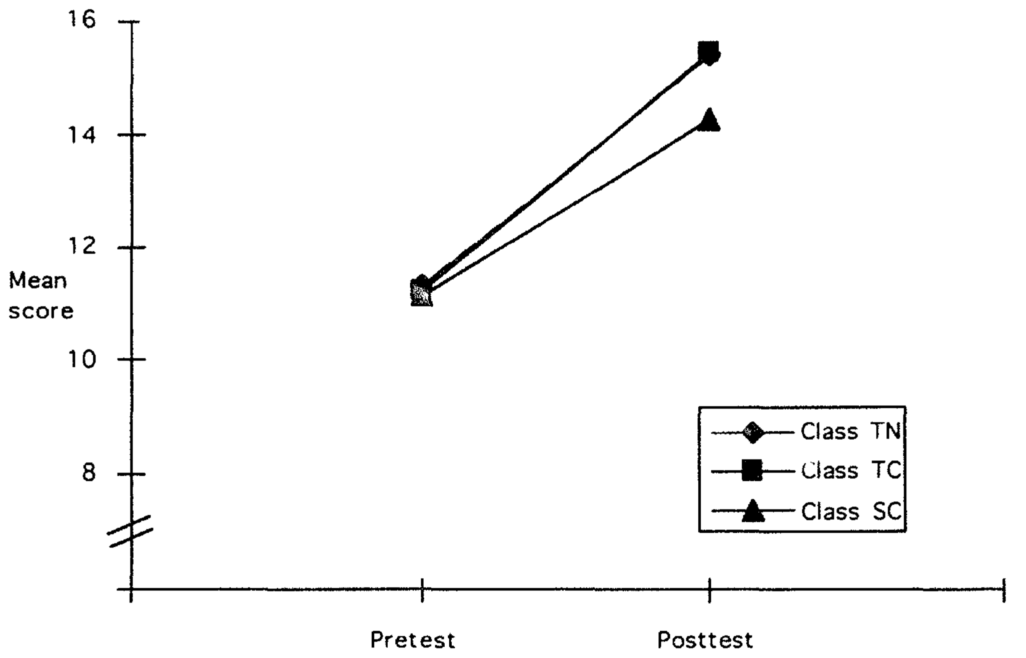


Figure 6. Mean total TOSIS pre and posttest scores for the three classes

TOSIS Scores for Specific Skill and Concept Areas: A Comparison of Classes on Pretest and Posttest Scores

The data were also analysed in terms of skill and concept areas. These subtests corresponded to identifying variables, writing an hypothesis; planning an investigation; drawing conclusions; identifying methodological limitations; and understanding concepts of hypothesis, theory, data and conclusions. The maximum possible score for these subtests was small (eg. 1 mark for *Writing an hypothesis*, 8 marks for *Planning an investigation)*, and because there was a limited number of items contributing to the subtest scores, it was decided that further statistical analysis would be inappropriate.

The descriptive statistics for each subtest are presented in Table 2 and the pretest and posttest means for each skill and concept area are represented graphically in a series of figures; Figure 7 (*Identifying variables*), Figure 8 (*Writing an hypothesis*), Figure 9 (*Planning an investigation*), Figure 10 (*Drawing conclusions*), Figure 11 (*Identifying methodological limitations*), and Figure 12 (*Understanding the concepts of hypothesis, theory, data and conclusions*). These data need to be interpreted with caution because some skill and concept areas are determined by one test item (Table 2) and because, for some items, small gains in the mean scores represent large a percentage increase. This is illustrated by the classes' mean gains and percentage increases for the skill *Writing an hypothesis*. *Identifying methodological limitations*, also based on one test item, had a minimum increase of 76% for Class SC. The relatively small gains in *Identifying variables* may be attributable to high pretest means and consequently a limited potential for gain.

In summary, although it was reported previously that there was no significant difference in the total TOSIS scores for the classes, graphs for pretest and posttest specific skill and concept areas show some differences. For the skills *Identifying variables* (Figure 7) and *Writing an hypothesis* (Figure 8) the improvements in performance for each class were similar. In contrast, for *Planning an investigation* (Figure 9) Class SC had slightly less improvement than the other classes, and for *Drawing conclusions* (Figure 10), *Identifying methodological limitations* (Figure 11) and *Understanding the concepts of hypothesis, theory, data and conclusions* (Figure 12), the improvements for Class SC were noticeably less than for the other classes.

Following Figures 7 - 12, the Chapter addresses students' competencies at planning investigations, conducting investigations, processing data and evaluating investigations from their performance of an investigation and indicated by students' written responses recorded on Investigation Planning and Report Sheets. In addition, the nature of students' responses is discussed.

Table 2. Mean pre and posttest scores, gain in mean score and percentage increase in mean score for skill and concept areas on the TOSIS for Classes TN (n = 23), TC (n = 22) and SC (n = 24)

Skill or concept area	Maximum score	Class	Pretest mean	Posttest mean	Gain in mean	Percentage increase
Identifying variables Q. 1, 6, 7 & 8	6	TN	4.87	5.48	0.61	13
		TC	5.23	5.64	0.41	8
		SC	4.79	5.38	0.59	12
Writing an hypothesis Q. 2	1	TN	0.35	0.91	0.56	160
		TC	0.32	0.91	0.59	184
		SC	0.25	0.83	0.58	232
Planning an investigation Q. 3	8	TN	2.48	4.22	1.74	70
		TC	2.18	3.77	1.59	73
		SC	2.42	3.75	1.33	55
Drawing conclusions Q. 4	2	TN	0.91	1.13	0.22	24
		TC	0.77	1.05	0.28	36
		SC	0.87	0.92	0.05	6
Identifying methodological limitations Q. 5	2	TN	0.39	0.87	0.48	123
		TC	0.23	0.86	0.53	274
		SC	0.33	0.58	0.25	76
Understanding the concepts of hypothesis, data theory and conclusions Q. 9-12	4	TN	2.30	2.83	0.53	23
		TC	2.50	3.23	0.73	29
		SC	2.46	2.83	0.37	15

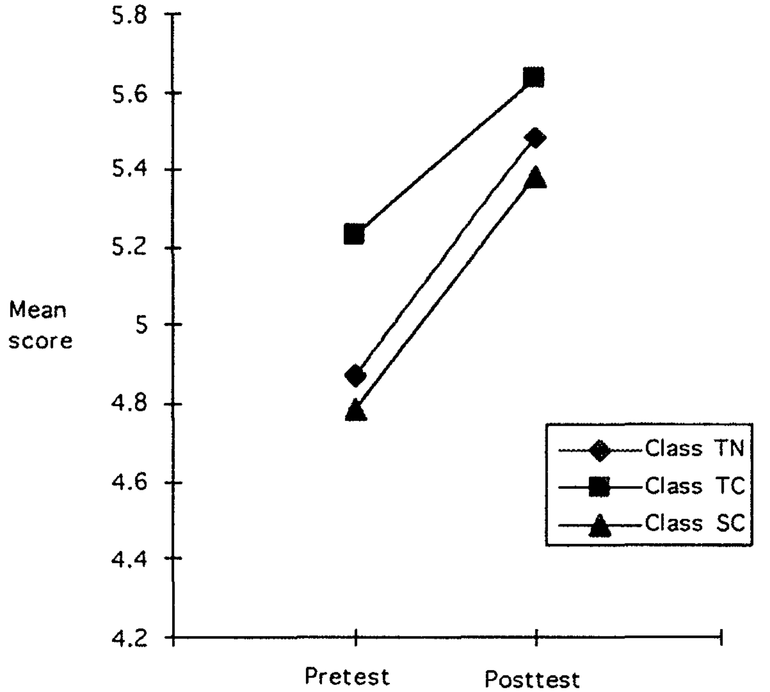


Figure 7. Identifying variables: Pretest and posttest TOSIS mean scores for the three classes

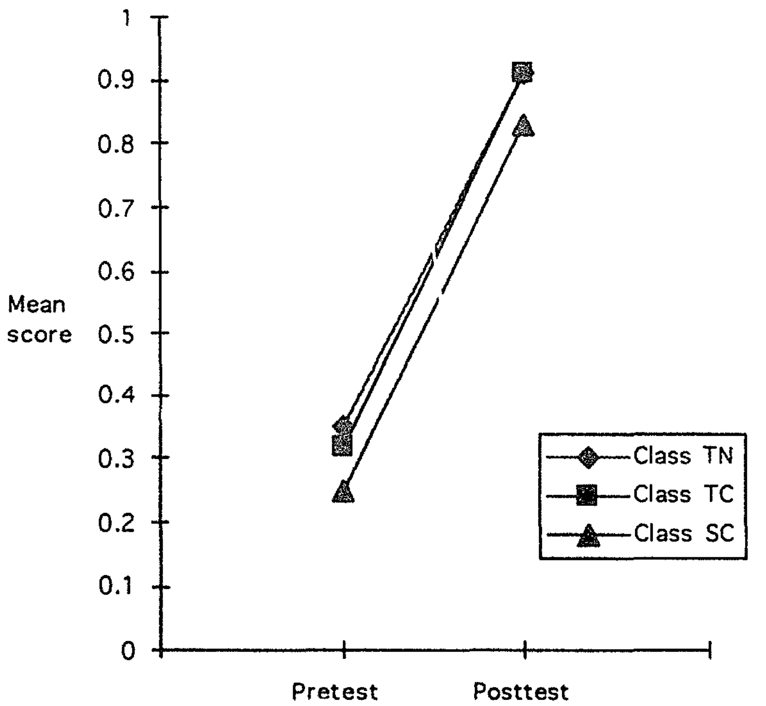


Figure 8. Writing an hypothesis: Pretest and posttest TOSIS mean scores for the three classes

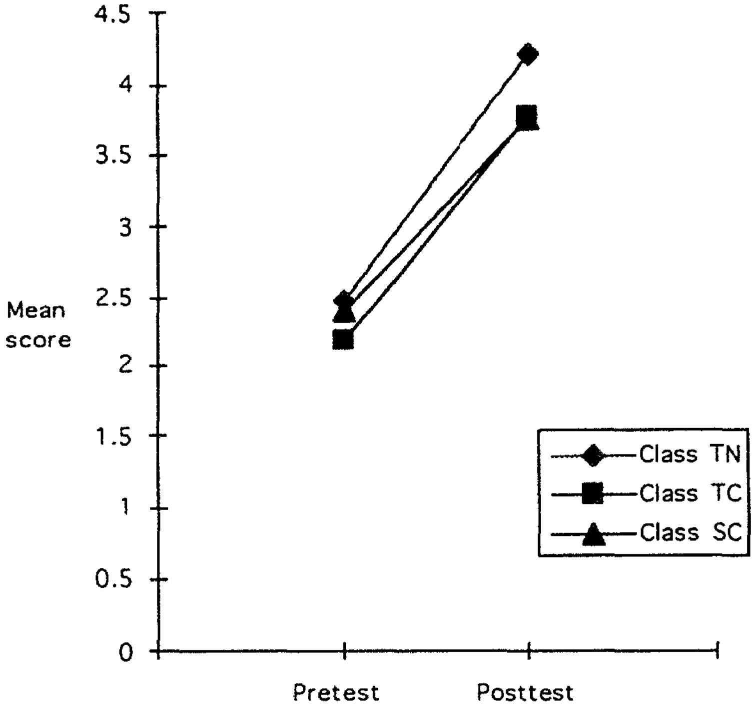


Figure 9. Planning an investigation: Pretest and posttest TOSIS mean scores for the three classes

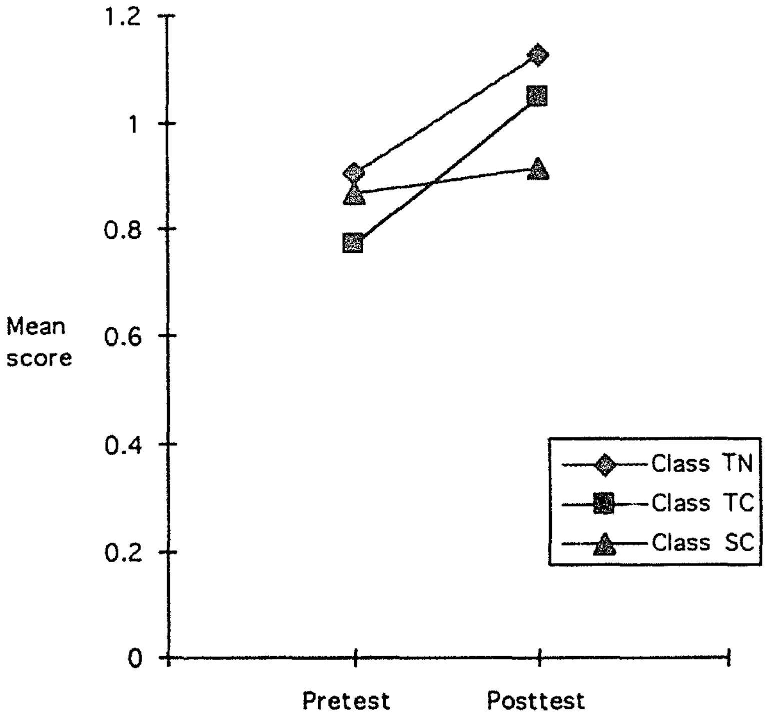


Figure 10. Drawing conclusions: Pretest and posttest TOSIS mean scores for the three classes

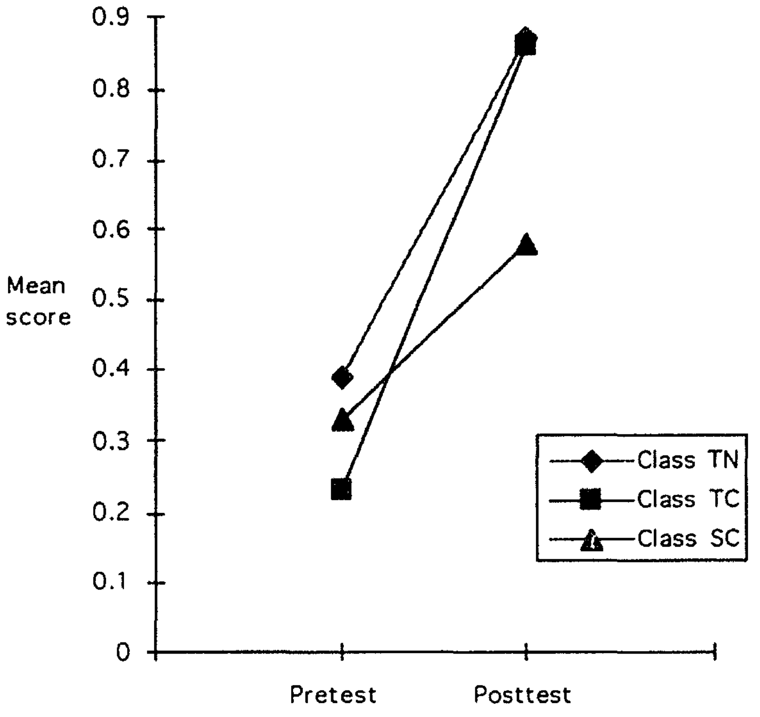


Figure 11. Identifying methodological limitations: Pretest and posttest TOSIS mean scores for the three classes

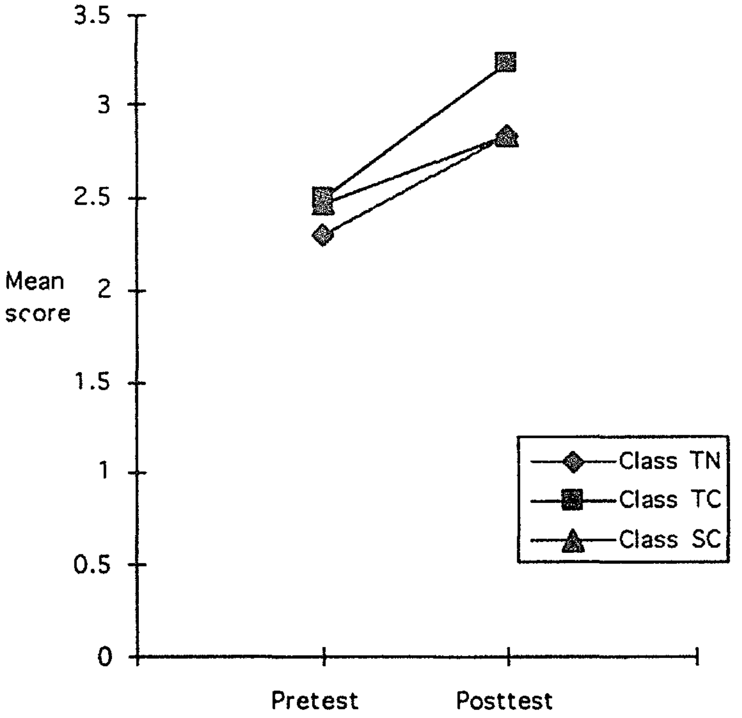


Figure 12. Understanding concepts of hypothesis, theory, data and conclusions: Pretest and posttest TOSIS mean scores for the three classes

Investigation Planning and Report Sheet (IPRS)

A Comparison of Classes on Pretest and Posttest Investigations

As previously described (Chapter 3) students' performances on the Investigation Planning and Report Sheets were coded into levels of achievement for each of the four substrands of an investigation; *Planning investigations, Conducting investigations, Processing data* and *Evaluating investigations.*

The number of students who achieved a particular level on the pretest and posttest investigation was tallied for each of the phases of investigating (Table 3). From these data it is evident that for each class there are improvements in students' performances from the pretest to the posttest. Students who failed to attempt a particular section of the IPRS are tallied as incomplete. For the pretest, these tallies represented an overall 76% completion for Class TN, 64% for Class TC and 72% for Class SC. Although all students completed the *Planning investigations* section of the pretest, the number of students failing to respond increased for the last sections of the IPRS. This may have been due the fact that students did not have an overview or conceptual model of the investigation. Their IPRSs suggested that they were disorganised and unable to finish in the allocated time, and/or that they did not understand what was required to complete the questions. This was evident from their IPRS which showed that many had collected insufficient or erroneous data for analysis.

Table 3: Number of students in the three classes who achieved at various levels for Planning investigations, Conducting investigations, Processing data and Evaluating investigations on the pretest and posttest

Level	Class TN (n = 21)		Class TC (n = 24)		Class SC (n = 20)	
	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
Planning investigations						
2	3	0	0	0	1	0
3	9	0	16	1	10	0
4	7	3	6	9	7	13
5	2	15	2	8	2	7
6	0	3	0	6	0	0
Incomplete[a]	0	0	0	0	0	0
Conducting investigations						
2	0	0	5	0	5	0
3	21	0	15	0	15	3
4	0	6	1	11	0	15
5	0	12	0	14	0	2
6	0	3	0	0	0	0
Incomplete	0	0	3	0	0	0
Processing data						
2	5	0	4	0	0	0
3	13	1	14	0	16	1
4	0	11	0	18	0	16
5	0	9	0	5	0	3
6	0	0	0	1	0	0
Incomplete	3	0	6	0	4	0
Evaluating investigations						
2	8	0	0	0	0	2
3	6	3	5	5	4	3
4	1	15	9	8	8	10
5	0	3	0	11	0	5
6	0	0	0	0	0	0
Incomplete	6	0	10	0	8	0

Note. [a] Incomplete refers to the number of students who did not completed the IPRS.

Students' performances on the posttest investigation are also summarised in Table 4. This shows the percentage of students in each class who achieved at Level 5. This Level is considered to be a benchmark for students in Year 9. It is evident that Class SC did not perform as well as Classes TN and TC for *Planning investigations, Conducting investigations* and *Processing data.* For *Evaluating investigations*, 25% of students in Class SC performed at Level 5, compared with 46% in Class TC and 14% in Class TN.

Table 4: Percentage of students for the three classes who achieved the benchmark of Level 5 on the posttest for Planning investigations, Conducting investigations, Processing data and Evaluating investigations

	Percentage of students achieving the benchmark standard of Level 5 on the posttest		
	Class TN	Class TC	Class SC
Planning Investigations	86	58	35
Conducting Investigations	71	58	10
Processing Data	43	25	15
Evaluating Investigations	14	46	25

The Wilcoxon Matched-Pairs Signed-Ranks test was used to compare the classes' pre and posttest levels of performance for the four phases of investigating. Details of the tests are summarised in Table 5 and indicate that in all cases the posttest levels were greater than the pretest levels, $p < .01$ for one tailed tests. The test therefore supports the descriptive data (Tables 3 & 4) and indicates that the improved performance due to test occasion was statistically significant.

Table 5. Analyses of changes from pretest to posttest score (levels) for the three classes on the four phases of investigation, using the Wilcoxon Matched-Pairs Signed-Ranks test

Class	Phase of investigation	Number of changes in level between the pretest and posttest			Z	p < .01 for a one-tailed test
		+ ve changes	- ve changes	ties		
TN	Planning	17	0	4	3.62	*
(n = 21)	Conducting	21	0	0	4.01	*
	Processing	20	0	1	3.91	*
	Evaluating	19	0	2	3.82	*
TC	Planning	22	0	2	4.10	*
(n = 24)	Conducting	24	0	0	4.28	*
	Processing	24	0	0	4.28	*
	Evaluating	21	3	0	3.77	*
SC	Planning	13	0	7	3.17	*
(n = 20)	Conducting	17	0	3	3.62	*
	Processing	19	0	1	3.82	*
	Evaluating	16	1	3	3.40	*

Level 5 represents an expected benchmark for performance of Year 9 students and Tables 3 and 4 indicate that fewer students in Class SC achieved Level 5 on *Planning investigations, Conducting investigations* and *Processing data*. Therefore additional analyses were conducted to determine whether differences between the classes performances were significant. The Kruskal-Wallis One-Way Analysis of Variance (Table 6) was used to determine whether the three classes were equivalent on the pretest measures of performance for the four phases of investigating. This test indicated that there was no significant difference ($p < .01$) in the classes' performances and confirmed the equivalence of the three classes at the pretest. However, on the posttest there was a significant difference between the

performance of the three classes for the planning phase (H = 9.76, df = 2, p < .01) and also for the conducting phase (H = 19.36, df = 2, p < .01). Therefore, these statistics confirm trends that emerged in the descriptive statistics for the IPRS (Tables 3 & 4) that showed fewer students in Class SC attained the benchmark level of performance on the posttest than students in Classes TN and TC. Also, they are consistent with results from the TOSIS which showed Class SC recorded smaller gains than the other classes for *Planning an investigation* (Figure 9), *Drawing conclusions* (Figure 10), *Identifying methodological limitations* (Figure 11) and *Understanding concepts of hypothesis, theory, data and conclusions* (Figure 12). Reasons for the more modest performance of Class SC on *Planning* and *Conducting investigations* are not evident from these data or from data gathered from the students' IPRSs. Class SC, however, experienced student assessed criterion referenced assessment whereas the other classes were teacher assessed and the poorer performance of the class may be due to the nature of feedback that students received after completing their investigations.

Table 6. Comparison of classes on pre and posttest performance in the four phases of investigation using the Kruskal-Wallis One-Way Analysis of Variance

Test occasion	Phase of investigation	H corrected for ties	df	p < .01
Pretest	Planning	0.2869	2	
	Conducting	6.3920	2	
	Processing	1.6948	2	
	Evaluating	1.9046	2	
Posttest	Planning	9.7627	2	*
	Conducting	19.3654	2	*
	Processing	3.3307	2	
	Evaluating	2.3823	2	

The Nature of Students' Responses to the IPRS

Planning investigations

During the pretest investigation it was intended that students would investigate factors that affect how quickly lipase breaks down fats in milk into fatty acids and 16% of students did this. Many students (66%) investigated factors that affect the acidity of milk, and for these investigations pH was the dependent variable instead of time. It is difficult to determine the reason/s why students did this but it may have been because they were provided with a colour chart for the pH changes of Universal indicator. A further 18% either did not complete the section or wrote hypotheses that could not be tested. The independent variables selected by students included the temperature, amount of milk, amount of lipase and type of milk.

Of the hypotheses that the students wrote on their IPRS, 42% were unlikely to be supported by data (Assertion 1). For example, 13% of students hypothesised, "The higher the temperature the more acidic milk becomes in the presence of lipase." This indicates that students applied the more general chemistry principle that an increase in temperature causes an increase in reaction rate and did not know that at high temperatures enzymes will denature and become ineffective (Assertion 2). Only one student (1.5%) investigated the converse of this hypothesis which is likely to be supported by data. Although in science it is quite common to propose an hypothesis that is ultimately not supported by data, Year 9 students in this study experienced difficulties when dealing with disconfirming data. The ways that students attended to the mismatch between their hypothesis and their data are discussed later in this Chapter.

Assertion 1: Many hypotheses that students investigate are unlikely to be supported by data.

Assertion 2: Many students think that an increase in temperature decreases the time taken for all reactions (including reactions involving biological enzymes).

It seems that the conventional wisdom 'the less you have the quicker it goes' and its converse 'the more you have the longer it takes to go' are reasonable rules that govern behaviour. However, in science these 'rules' only apply in some instances. In chemistry they apply when reactions go to completion. The 'rules' do not apply when referring to rates of reaction. For example, the time taken for the **completion** of a reaction between ions in concentrated solutions is **longer** than for ions in dilute solutions of the same substance. In contrast, the initial **rate of change** for concentrated solutions is **greater** than for dilute solutions. Hence, if students formulate their hypothesis with the expectation that they are measuring the time taken for the reaction to go to completion and actually measure the rate of the reaction then they are likely to draw conclusions that are inconsistent with established science. For example they may collect data for a range of concentrations over a predetermined period of time or until they obtain a predetermined amount of product. In these instances they will measure the rate of the reaction. Hence, if they assume that they are measuring the completion of the reaction then they are likely to conclude that the more concentrated solution reaches completion before the less concentrated solution. A specific example follows and these ideas are expressed as Assertion 3.

Some students (12%) varied the type of milk and investigated the hypothesis, "The more fat in the milk the longer the milk takes to become acidic in the presence of lipase." It is likely that the students believed that with more fat in solution the reaction would take longer to reach completion. Using the rates of reaction concept, solutions with more fat molecules will initially react with lipase more quickly and

the time taken for the solution to become acidic will be less, although it will take longer to reach completion.

Assertion 3: Some students interpret the time taken for the reaction to mean the time for the reaction to go to completion. This results in difficulties when they collect their data over a fixed time period and hence measure the rate of the reaction.

Another problem for students arose when they did not keep the volume of the reaction constant. A number of students (17.5%) investigated how the amount of milk affects the acidity of the resulting solution and hypothesised, "The less milk (fat) the quicker it becomes acidic in the presence of lipase." These hypotheses were confirmed by students' data because they had not kept the reacting volume constant and the concentration of the enzyme was higher in a smaller volume of milk. However, the hypothesis is unlikely to be supported if the volume of the reacting substances is kept constant because with a more dilute solution the reaction rate is slower. These students did not have an appreciation that volume must be kept constant for a fair test (Assertion 4).

Assertion 4: Some students who investigate the effect of the amount of substance in a chemical reaction on the time taken for the reaction, do not control for the volume of solution.

For the pretest most students indicated that they understood the notion of a fair test and the need to keep some factors constant and were assigned Level 3 according to the Student Outcome Statements (Appendix A; Education Department of Western Australia, 1997). Students found the identification of the variable to be changed and the dependent variable to be measured (Level 4) difficult in the pretest. The dependent variable could be measured by recording the time taken for the

Universal indicator to undergo a specific colour change in response to changes to the independent variable as outlined in Figure 5. As mentioned previously, many students investigated factors affecting the amount of fatty acids produced by milk and they found it extremely difficult to determine how to measure the dependent variable because there was virtually only one colour change for the Universal indicator. This procedure would be acceptable if students recorded the colour change over a fixed period of time. However, almost all of these students showed no evidence of timing the reaction.

For the posttest investigation 35% of students investigated hypotheses that were unlikely to be supported by the data. Twenty three percent of students hypothesised the higher the temperature the quicker the trypsin reacts with protein; 4.5% hypothesised the more basic the solution the faster the reaction; 3% hypothesised the less basic the faster the reaction; and 4.5% the more solution (with a fixed trypsin to solvent ratio) the faster the trypsin will react.

More students wrote an hypothesis and were able to use the hypothesis to focus their planning. They were able to list the variables important in the data collection and planned to control several of them. In addition, they planned data collection procedures, equipment and techniques to be used (Level 5). For example, Kit hypothesised, "That the trypsin reacts slower as the range between the testing temperature and 37 °C increases," and her plan revealed a sound understanding of the variables that she planned to control and how she would measure the variables. She stated that her preliminary trials at 37 °C and 23 °C were used to see if the data collection were feasible. However, after the investigation she realised that her planning was not adequate because she tested over too large a range of temperatures (between 10 °C and 84 °C) and the trypsin took too long to react.

Kit: (Referring to the preliminary trials) We simply timed how long it took for the protein to react with the trypsin (for the film to go clear) at 37 °C and 23 °C. As these results seemed reasonable, we began our investigation.

> However, if we had more time we would have realised that such an extensive temperature range caused many problems (eg. took too long) and so we may have changed our plan.

Most students used preliminary trials to observe the phenomenon being investigated, and to organise their data collection (Assertion 5). Most did not use the preliminary trials to determine the range over which they should collect data or the intervals of data collection (Assertion 6). Typical of students' responses are those by Pippa, Kelly, Kim, Emma and Tina and these comments substantiate Assertions 5 and 6. They wrote how they would observe the phenomenon (Pippa) and organise general aspects of the investigations such as the number of trials (Kelly); the management of time (Kim); the water bath and how to measure the dependent variable (Emma); and how to organise control of the temperature (Tina).

Pippa: We trialed it to see how it would work.

Kelly: We changed from doing 4 experiments to 3. We also changed the amount of water we were going to use (in the water bath).

Kim: We decided to put a couple of test tubes into the water bath at the same time because doing it seperate (sic) (1 at a time) it took too long.

Emma: By doing these trials we found that the tap water was better experimented in a big beaker (water bath), also it was hard to be accurate with the clearence (sic) in the proten (sic), so we waited till it was completly (sic) clear and then stopped the timer.

Tina: Regarding the 70 °C water we attempted to try and "top up" the beaker with more 70 °C water. It didn't keep the temperature constant but merely slowed the rate of heat loss. We also increased our monitoring of the film in order to get an acurate (sic) time.

Assertion 5: When planning investigations most students use preliminary trials to observe the reaction/experiment and to organise procedures.

Assertion 6: When planning investigations most students do not consider the range over which the data should be collected or the data collection interval.

For the posttest investigation a few students achieved Level 6 because they were able to analyse the problem; formulate a question or hypothesis for testing; use scientific knowledge to identify the main variables to be considered and make predictions; and plan for accurate measurements. For example, Vaille proposed to test the hypothesis, "The higher the temperature the faster the trypsin breaks down the gelatin into protein," and, after a highly detailed description of her plan, she outlined her preliminary trials and the subsequent changes that she made. Vanessa also achieved Level 6 for *Planning investigations* because she explained that she set out to operationalise the measurement of the dependent variable.

Vaille: The first preliminary trial that we did was trying to increase the temperature each time by 2 °C. We started from 35 °C and increased the temperature by 2 °C up to 45 °C. This experiment did not work out because it took a long time and there were too many to do (10 experiments from 25 °C to 45 °C). We started from 35 °C, 37 °C, 39 °C to 41 °C and we found out that 2 °C difference was not that much and we could not tell the difference, as the time was only 2 seconds apart. As we realised that it did not quite work out we increased the temperature and did another preliminary trial. With the second preliminary trials, the experiment seemed to be going pretty well. As a result of our first preliminary trials we changed the temperature increase each time from 2 °C to 5 °C. With a difference of 5 °C each time it's not time consuming and the experiment just turned out to be what we expected.

Vanessa: We also need to do trials to work out when the trypsin solution has broken down the gelatin. (Later, after the trials she adds…) We are going to record our results from when the film turns completely clear.

Conducting investigations

Assessing students' abilities to conduct investigations was based on the information presented in this section of their IPRS. From the pretest to the posttest investigation, students improved in the amount of detail presented. As well, they made more references to organisational factors, including group management (Rose) and time management (Gill). This is supported by questionnaire data (Chapter 5) in which they identify working cooperatively, managing time and being organised, as investigations competencies that they have learned (Assertion 22, p. 121).

Rose: One way of making sure our data is accurate was to make sure everybody did the same things and each temperature was the same.

Gill: Working quickly we will set up four beakers ...

There was a tendency amongst some students to observe changes that they thought would happen rather than changes that actually occurred. This is supported by teacher interviews with Mrs Grant (p. 153) and student interviews with Group TC (p. 176). It is also exemplified on Marilyn's IPRS and on the IPRS of other members in her group, Dolly and Kara. At eight minutes they wanted Universal indicator to change colour in the lipase and milk solution while at 14 minutes they wanted the colour to stay the same. Dolly and Kara wrote similar comments to Marilyn.

Marilyn: 8 mins - Both (the reactions in) the water bath and normal water looked as if they are **trying to change** (colour).

14 mins - **No change** except the hot bath **has changed** colour to darker orange.

Assertion 7 is supported by previous research findings. Researchers (Duggan & Gott 1995; Foulds, Gott & Feasy, 1992; Germann & Aram, 1996) report that students frequently interpret their results on the basis of what they think will happen rather than the data before them. Furthermore, students' beliefs are held with certainty (Kuhn, 1992) and are persistent and resistant to change (Novak, 1988; Novick & Nussbaum, 1982).

Assertion 7: Some students 'observe' changes that they think will occur in investigations rather than changes that actually occur.

Overall, from the pretest to the posttest there seemed to be a greater appreciation of the need to control variables in an effort to make sure the data were not influenced by external factors (Level 4). In addition, students seemed to have

developed an appreciation of the need to replicate or repeat the data collection, average the results, and process the averaged results (Level 4). These ideas are expressed in Assertion 8.

Assertion 8: Most students learn to repeat or replicate the data collection and to average the results.

When students averaged their results, several assumed that time was a decimal measure (Assertion 9). For example, Pearl wrote down the following times from her stopwatch, 1 55.02; 1 59.56 and 2 10) and averaged them to arrive at 1.75 minutes instead of 1 minute 58.22 seconds. Similarly, Heather reported an average time of 1 minute 75 seconds. In addition, it was evident that students had not developed an appreciation of the degree of accuracy with which readings should be made (Assertion 10) and this is exemplified by Pearl's records of the disappearance of trypsin to points of a second accuracy. This could be expected because students are typically not taught about significant figures and the uncertainty of measurement until later high school years.

Assertion 9: Some students treat time as a decimal measure and consider seconds as one-hundredths of a minute.

Assertion 10: Most students do not have an appreciation of the degree of accuracy to which measurements should be recorded.

There was some overlap between Level 5, *Conducting investigations* and Level 6, *Planning investigations*. In this study students who planned for accurate data measurements (Level 6, *Planning investigations)* also conducted preliminary trials of the experimental procedure to improve the procedure and/or measurement techniques (Level 5, *Conducting investigations*). Therefore, students who achieved

Level 6 for *Planning investigations* also achieved Level 5 for *Conducting investigations*. Students were unable to achieve Level 6 for *Conducting investigations* because they were not provided with the opportunity to decide what equipment was needed and request equipment for the investigation. Some students however, did achieve some of the pointers described by the Student Outcome Statements such as using operational definitions to enhance the consistency of the measurement and recognising the difficulties in making accurate measurements.

Processing data

For the pretest many students gathered insufficient data to present in tables, draw graphs and identify patterns and, as a result, they were unable to summarise patterns in the data (Level 3). For the posttest, all students completed all sections of the IPRS. Students also collected more data and almost all had conducted repeated trials and averaged their data. Most plotted line graphs where appropriate and consequently they were able to make conclusions by summarising and explaining patterns in the data (Level 4). Some students made conclusions that were consistent with their data and, in addition, gave reasons for their conclusions based on scientific knowledge (Level 5). To achieve Level 5 requires a description of patterns in the data and an explanation of those patterns in terms of scientific knowledge. For this investigation students were expected to explain the phenomenon being studied in scientific terms rather than illustrate an understanding of reaction rates because this concept had not been studied. In addition, Level 5 required students to relate one phase of the investigation to another so that "Conclusions are consistent with the data..." (Education Department of Western Australia, 1997). This ability to relate different aspects of investigating requires hypothetico-deductive reasoning (Lawson, 1995) or metacognition (Kuhn, Amsel & O'Loughlin, 1988) and supports the notion that the whole investigation is greater than the sum of the parts (Hodson, 1992). More specifically Foulds et al. (1992) report that many students fail to see the importance of evidence in drawing conclusions. Tina's response to the

processing data section of the investigation is typical of Level 5 achievement. She tested the reaction rate of trypsin with gelatin at 6, 22, 36, 60 and 70 °C; and stated the following:

Tina: In conclusion our hypothesis was supported. We also concluded that the further the temperature is away from body temp. whether it be colder or hotter, the longer the trypsin will take to react with protein (eg. 70 °C and 6 °C took over 20 minutes longer to react than body temperature).

The *Processing data* section of the IPRS was characterised by a wide range of students' responses. It appeared that many students had a confirmatory bias and wanted to make a positive finding. As previously discussed, some students seemed to unintentionally let this bias cloud their interpretations; a phenomenon previously reported in the research literature (Duggan & Gott, 1995; Foulds, et al., 1992; Germann & Aram, 1996). Other students seemed to intentionally change or modify the hypothesis. This occurred when the hypothesis they chose was unlikely to be supported if they 'correctly' interpreted their data. Clearly from the students' perspective there were two solutions to this dilemma; either change the hypothesis or change the results. In science it is unethical to change data but it is less unethical to change an hypothesis or to finalise research questions after the research has been conducted. It is likely that students are unaware of these ethical issues. Hence, for the student, it may be that they perceive changing the data equally acceptable to changing the hypothesis. Either option would provide them with a sense of 'getting things right' and it is arguable that previous successes at school were based on correct answers (Fairbrother & Hackling, 1997). The notion of "fudging" experiments to get the "right answer" has been previously reported in the literature (Rigano & Ritchie, 1995, p. 359).

The following discussion deals with ways that students confirmed their hypotheses (Assertions 11, 12, 13, 14 and 15). Some students, however, obtained confirmatory data to a scientifically correct hypothesis and their responses to the

Processing data section of the IPRS are not discussed. In addition, a few students rejected their hypothesis (Assertion 16), and some had problems dealing with continuous data (Assertion 17).

Some students **confirmed their hypothesis with insufficient data** (Assertion 11). For example, Christie, Jody and Harue hypothesised, "The more fat in the milk the higher the acidic rating." They conducted a test on each of two low fat milk samples and one whole milk sample to discover that one of the low fat samples and the whole sample had pH readings of between 5 and 6, and that the second low fat milk sample had a pH of between 6 and 7. From these limited data, of which one set of data was contrary to their hypothesis, they accepted the hypothesis stating, "The more fat in the milk the more acidic the rating when passed through the small intesstine (sic)," and added "Our test results tell us this." Similarly, with insufficient data Vaille, Mary and Naire accepted their hypothesis, "The more lipase you put into the milk the more fatty acid you will get." Their data collection revealed that they only conducted one test with lipase. In addition, for the trypsin investigation Gill, Rose and Pippa confirmed their hypothesis that "Gelatin reacts with trypsin best at 37 °C" with data at three different temperatures none of which was 37 °C. Their tests were conducted at "room temp (22 °C), freezing (8 °C), oody temp (41-42 °C), and boiling (70-80 °C)." Clearly they should have collected more data around 37 °C or modified their hypothesis.

Assertion 11: Some students confirm their hypothesis with insufficient data.

Some students **confirmed their hypothesis with highly dubious data** (Assertion 12). For example in the trypsin investigation Jo, Ed and Janet hypothesised, "That the higher the temperature the faster the reaction will occur." They collected data for boiling water, tap water (24 °C) and icy water with the temperature of the boiling water and the icy water unspecified. They averaged three

sets of data for each temperature to obtain times of 3.39 s (boiling water), 11.26 s (tap water) and 55.72 s (icy water). Clearly the result for boiling water are not accurate as the enzyme would have denatured at high temperatures and the students either have fudged their results or have incorrect temperature readings. This is consistent with research by Rigano and Ritchie (1995).

Assertion 12: Some students confirm their hypothesis with data from dubious sources.

Some students **confirmed their hypothesis by ignoring or rejecting data** (Assertion 13). For example in the trypsin investigation, Marion and Beth found the average times for the rate of reaction at specific temperatures to be 13.19 at 21 °C, 5.15 at 37 °C, 0.20 at 40 °C, and 10.22 at 62 °C. No units were stated for the time. The students, however, confirmed their hypothesis stating, "The higher the temperature the faster the reaction/the time taken." Marion wrote that she arrived at this conclusion "because as the temp. increased, the reaction/the time taken is faster." In a similar vein, Susan and Vanessa confirmed their hypothesis, "The larger the amount of fat in milk the slower the lipase will take to break down into fatty acids." Their data indicated the opposite although they wrongly plotted two points on their graph and may have reached this conclusion from the graphed results rather than the results in the table.

A major difficulty for some students resulted from their understanding of the relationship between temperature and reaction rate. The problem surfaced because some students chose to investigate the effect of temperature on chemical reactions, as in the catalyst investigation and the Panadol investigation, and from these investigations they concluded, "The higher the temperature the faster the reaction rate." When they performed the trypsin investigation some of these students decided again to investigate the affect of temperature. This time they found that the reaction

did not work at high temperatures. Many of these students disregarded the high temperature results and dismissed the result saying that it didn't work. For example the group comprising Janet, Jo and Ed dismissed the result by attributing the problem to their inability to control the temperature of the water bath. This notion addressing the rejection of significant data to confirm the hypothesis is expressed in Assertion 13.

Janet: When we did the boiling water we had problems keeping the temperature in one place so we disregarded it and included the water bath instead (40 °C).

Assertion 13: Some students ignore or reject data so that they can confirm their hypothesis.

Some students **rejected their hypothesis and formulated a new conclusion based on insufficient evidence** (Assertion 14). For example, during the lipase investigation Gay, Kim and Tammy initially hypothesised, "The higher the temperature of the water the more acidic the milk becomes in the presence of lipase." They conducted the investigation at 4, 43 and 80 °C. They found that at 43 °C the lipase produced a milk solution of pH 5 while at other temperatures the solution remained at pH 6. They ignored the result at 43 °C and rejected their original hypothesis. Then they concluded that, "The temperature doesn't affect the pH level of milk in the presence of lipase," and added, "The results of our investigation proves this." Clearly the students were correct in rejecting their original hypothesis, however, the conclusion that they reached was based on insufficient evidence. If they had used their limited data to form a conclusion then they would have said that, "Milk becomes more acidic in the presence of lipase around 43 °C."

Assertion 14: When faced with unexpected results some students reject their original hypothesis and formulate a new conclusion based on inconclusive evidence.

Some students **confirmed a different hypothesis from that which they initially planned** (Assertion 15). For example, Kit and Sally hypothesised, "If the amount of lipase is increased, then the rate of reaction will increase." In one test tube they placed 1 mL of lipase and in the other 5 mL. Then, instead of measuring the time for the reaction they monitored the pH by looking for colour changes with Universal indicator. In their conclusion they state that, "The amount of lipase does affect the pH value of the experiment once the reaction takes place." They made no reference to the reaction time and did not collect any data for the time of the reaction.

Assertion 15: Some students collect data in order to confirm a different hypothesis from that which they originally planned to investigate.

Some students **investigate an hypothesis that cannot be supported and use their own data to reject the hypothesis** (Assertion 16). This was encouraging because from the preceding discussion, it is evident that most students in this study had difficulties when dealing with unexpected or disconfirming data. Perhaps the majority of students do not realise that it is scientifically acceptable to reject an hypothesis based on disconfirming evidence. On the other hand it may be postulated that the intellectual honesty required to deal with unexpected data is rare in Year 9 students. Encouragingly, some students (Rhyll, Simone and Sarah) said that they did not confirm their hypothesis. They performed a carefully planned investigation to test the hypothesise that, "The more solvent you use for the reaction the faster the reaction of breaking down gelatin will be." After conducting three trials with different volumes of solution and the same solute to solvent ratio, the students

acknowledged that they had found virtually no difference in the data. Simone wrote " ... our hypothesis was proven wrong," and added, "I think our conclusion was the most appropriate for the amount of trials that we did."

Assertion 16: Some students investigate an hypothesis that cannot be supported and use their data to reject the hypothesis.

Some students **viewed continuous data as discrete data** (Assertion 17). For example Gill, Rose and Pippa used a column graph to plot the time for the reaction versus the amount of milk. Similarly Alice, Stephanie and Lara used a column graph to plot pH versus the amount of milk. Nicki, Kerry and Diana failed to use a scale on the temperature axis (horizontal axis) but did so on the time axis (vertical axis). They effectively drew a column graph for temperatures of 6°C, 20 °C etc. with each column approximately 2 mm apart. The notion of presenting temperature as a discrete variable was evident in the work of other students. Some of these students plotted column graphs with the temperature categories, cold/freezing, warm, hot and very hot in random order on the horizontal axis. It is likely that some students who plotted continuous data as discrete data viewed the data from a comparative perspective rather than as contributing to a trend or pattern. This was particularly evident when students did not average or collate the results in any way, but graphed each set of data as a separate incident. For example, nine students plotted each pH reading they made during the lipase investigation. Some used two column graphs to illustrate 'before' and 'after' scenarios and others plotted one column graph that showed both 'before' and 'after' columns. These findings are consistent with students' difficulties associated with graphing previously identified by Roth and McGinn (1997).

Assertion 17: Some students view continuous data as discrete data.

Evaluating investigations

Many students identified difficulties that they experienced in doing the investigations (Level 3). For example, Tina implied that it was difficult to maintain a constant temperature for the reaction and stated that the main source of error was, "The testing (of trypsin and photographic film) in 70 °C water as we could not stabilise the temp." At Level 4 students made general suggestions for improving the investigation. Their comments were about needing better equipment, making more exact measurements and doing more testing. Mary's suggestion about improving the control of variables typifies a Level 4 response.

Mary: If we were to do this again I would like to divide the jobs up for each person (1 person for the hot, 1 for the warm etc.)

To achieve Level 5, students were required to suggest specific changes that would improve the techniques used or the design of the investigation. This could involve saying how to improve the accuracy of the measurement procedure, identifying a variable that was not kept the same across treatments and saying how this could be achieved, or saying how the measurement procedure could have been applied more consistently. Sarah's comments typify a Level 5 response.

Sarah: The main errors were in taking the film out every 10 seconds. It was hard to be exact. I think our results were wrong. Doing it again I would have used trypsin from the same batch and kept them all the same temperature.

Few students' evaluations indicated that they were achieving Level 6. At this level students are required to recognise inconsistencies in the data, identify the main source of error, and to suggest improvements that would reduce the sources of error. Students such as Penny varied the amount of trypsin between 2 mL and 8 mL in the reaction with photographic film. The difference in the time for the reaction of the gelatin was small so she suggested that their group should use a wider range of amounts of trypsin. Other students who only conducted a few tests over a very wide

range of temperatures suggested that they would repeat the investigation using a smaller temperature range.

One of the dilemmas in *Evaluating investigations* is that the better students are at planning and conducting investigations and processing data the less opportunity they have to make suggestions to improve the investigation. For example, in the preceding paragraph Penny achieved Level 6 for *Evaluating investigations* because, after completing the investigation, she realised that she should have gathered data over a wider range. In contrast, Vanessa and Susan realised that this was important at the outset of the investigation and conducted preliminary trials to identify the range over which they would gather their data and although they received Level 6 for *Planning investigations* and *Conducting investigations* they received Level 5 for *Evaluating investigations*. This anomaly with assigning Levels of achievement to the *Evaluating investigations* substrand is exemplified by the pretest and posttest scripts of Diana and Helen. In her pretest Diana states that she has not controlled all the variables and adds that she would be more careful with the quantities she used. Hence, she was assigned Level 4 because she made "general suggestions for improving the investigation." In contrast, in her posttest she says that she has been more accurate and, as a result, does not make general suggestions to improve the investigation. Consequently she was not assigned Level 4 in her posttest investigation. Similarly, Helen makes general suggestions for improving her pretest investigation (Level 4), but does not do this for her posttest investigation.

Diana: (Pretest) The quantities of indicator and sodium (sic) used may have varied however we are confident our conclusion is correct. In repeating this experiment we would be more careful of the quantities used and we would set ourselves more time. (Level 4)

Diana: (Posttest) We found no main sources of error in our experiment. We are confident that the conclutions (sic) are accurate. We would allow more time to trial and study the experiment. (Level 3)

Helen: (Pretest) The main sorces (sic) of error in our experiment is that we forgot to time how long it took for the milks to change colour in the water bath and this should of determined our results. We are not very confident of our conclusion although if our hipothesis (sic) was "The amount of acid in different milks' (sic)" our confidence would be greater. We would time how long it took for the milk to change colour and rewrite the method. (Level 4)

Helen: (Posttest) Our main source of error was the accuracy of our results. Although our hypothesis was supported by our results we could have been more precise. We are confident that our results are supported although not to (sic) much detail. (Level 3)

When asked, "If you were to do this investigation again how would you change the investigation?" common responses were that students would change the hypothesis. Kim's comment below, is consistent with the previous discussion about the dilemmas students experience when they realise that their hypothesis is not scientifically correct. The perception that most students believe they need to test a scientifically accurate hypothesis is expressed in Assertion 18.

Kim: Our hypothesis was not right. ...start off with a new hypothesis.

Assertion 18: Most students perceive that it is necessary to investigate and confirm a scientifically 'correct' hypothesis.

Some students, Carol and Simone comment that they would investigate different variables if they were to repeat the posttest. This may be due to the fact that they were unable to confirm their hypothesis with the data that they had collected.

Carol: If we did it again we would probably change the temperature instead of the amount of sodium hydroxide.

Simone: If I was to do this again I think I would choose another variable to investigate that would hopefully be more consistent.

Summary of the Chapter

This Chapter compared students' pretest and posttest investigation competencies to gauge the effect of the instructional program. Two measures of investigation competencies were used; a pencil and paper Test of Science Investigation Skills (TOSIS); and a practical investigation that students reported on an Investigation Planning and Report Sheet (IPRS) as they performed the investigations. In addition, qualitative data gathered from students' responses to the IPRS were discussed, and 18 assertions about their investigation competencies were proposed.

The descriptive statistics relating to the TOSIS indicated that gains in mean total test scores were achieved by each class; between 3.16 and 4.13 out of a total score of 23 (Table 1). A two-way ANOVA was conducted and showed that there was a significant effect for test occasion but no significant difference between the classes and no interaction effect. Descriptive statistics for subtests that comprise the TOSIS were presented (Table 2). The biggest gains in performance were for *Writing hypotheses* and for *Identifying methodological limitations*. Gains were also recorded for the subtests, *Identifying variables, Planning an investigation, Drawing conclusions,* and *Understanding the nature of hypothesis, data, theory and conclusions*. Of the three classes, Class SC had the most modest gains for *Planning an investigation, Drawing conclusions,* and *Understanding the nature of hypothesis, data, theory and conclusions*.

The IPRS was used to collect information about students' investigation competencies of *Planning investigations, Conducting investigations, Processing data and Evaluating investigations*. Students' performances were judged according to standards of achievement described in the Student Outcome Statements (Education Department of Western Australia, 1997) and students were assigned

levels of achievement. The Wilcoxon Matched-Pairs Signed-Ranks test compared pre and posttest levels and indicated that the posttest levels were significantly greater than the pretest levels for all classes, $p < .01$ for one-tailed tests (Table 5). The Kruskal-Wallis One-Way Analysis of Variance indicated that on the pretest investigation there was no significant difference between the classes, however, on the posttest investigation there was a significant difference between classes for *Planning investigations* ($H = 9.76$, $df = 2$, $p < .01$) and *Conducting investigations* ($H = 19.36$, $df = 2$, $p < .01$) (Table 6). On the posttest investigation, fewer students in Class SC achieved the benchmark standard of Level 5 for *Planning investigations* and *Conducting investigations* than in Classes TN and TC (Table 4). Class SC also made more modest gains on substests of the TOSIS when compared with Classes TN and TC.

Eighteen assertions (Figure 13) were formulated about students' investigation competencies from the scoring of the IPRS and these have been presented under the general themes of *Planning investigations*, (Assertions 1 to 6), *Conducting investigations* (Assertions 7 to 10), *Processing data*, (Assertions 11 to 17) and *Evaluating investigations* (Assertion 18). Four assertions associated with *Processing data* related to fudging data. Assertions 11, 12 and 13 focus on fraudulent practices used to confirm the student's original hypothesis. Assertion 14 relates to students rejecting their original hypothesis and deciding upon a different conclusion, regardless of the conclusiveness of their data. These practices appear to have been adopted by students to get the right answer, and have been previously described by Rigano and Ritchie (1995).

The following Chapter discusses the data gathered from the student questionnaires.

Assertions

Planning investigations

1 Many hypotheses that students investigate are unlikely to be supported by data.

2 Many students think that an increase in temperature decreases the time taken for all reactions (including reactions involving biological enzymes).

3 Some students interpret the time taken for the reaction to mean the time for the reaction to go to completion. This results in difficulties when they collect their data over a fixed time period and hence measure the rate of the reaction.

4 Some students who investigate the effect of the amount of substance in a chemical reaction on the time taken for the reaction, do not control for the volume of solution.

5 When planning investigations most students use preliminary trials to observe the reaction/experiment and to organise procedures.

6 When planning investigations most students do not consider the range over which the data should be collected or the data collection interval.

Conducting investigations

7 Some students 'observe' changes that they think will occur in investigations rather than changes that actually occur.

8 Most students learn to repeat or replicate the data collection and to average the results.

9 Some students treat time as a decimal measure and consider seconds as one-hundredths of a minute.

10 Most students do not have an appreciation of the degree of accuracy to which measurements should be recorded.

Processing data

11 Some students confirm their hypothesis with insufficient data.

12 Some students confirm their hypothesis with data from dubious sources.

13 Some students ignore or reject data so that they can confirm their hypothesis.

14 When faced with unexpected results some students reject their original hypothesis and formulate a new conclusion based on inconclusive evidence.

15 Some students collect data in order to confirm a different hypothesis from that which they originally planned to investigate.

16 Some students investigate an hypothesis that cannot be supported and use their data to reject the hypothesis.

17 Some students view continuous data as discrete data.

Evaluating investigations

18 Most students perceive that it is necessary to investigate and confirm a scientifically 'correct' hypothesis.

Figure 13. Summary of assertions from scoring the Investigation Planning and Report Sheet

CHAPTER 5

STUDENT RESPONSES TO THREE QUESTIONNAIRES

Overview of the Chapter

This Chapter discusses students' responses to three questionnaires. Students completed the questionnaires after each of three treatments of 10 lessons during May, August and November. The questionnaires differed on each occasion because the investigation tasks were different, and because the classes had different assessment regimes. The sequence of tasks was described in Chapter 3 and is outlined in Appendix B. The investigations for Class TN were assessed by the teacher and graded on a norm referenced scale, Class TC was teacher assessed and graded according to a set of criteria, and Class SC was student assessed and criterion referenced. Condensed sample copies of the questionnaires for Class TN (May) Class TC (August) and Class SC (November) are presented in Appendix M.

Students' perceptions of the investigation competencies, the cognitive apprenticeship model of instruction and the assessment regimes are addressed in the Chapter. First, in relation to developing investigation competencies, students' perceptions of the amount of learning from different tasks in terms of science content knowledge and procedural knowledge are presented. In addition, students' perceptions of the investigation competencies they learned are categorised, examples of their responses for each category are presented, and the competencies addressed by each investigation are highlighted. In these analyses students' responses to the same questions were pooled across classes and over time.

Second, students' perceptions of the cognitive apprenticeship model of instruction are addressed. This section begins with a discussion of students'

perceptions of the **best** way to learn about doing investigations. By ranking preferred ways of learning, comparisons can be made about the relative worth of the following in terms of learning about doing investigations; students performing investigations in a group and by themselves, articulating with the teacher and with their peers about investigations, teacher modelling of investigations, and teacher and student marking/correcting investigations. These data are used to gain insights into specific strategies associated with the cognitive apprenticeship model of instruction such as articulating and coaching, and teacher modelled investigations. Students' perceptions of teacher modelled investigations are also accessed from questions in which students identified the task that made the greatest contribution to their learning about the phases of investigating; *Planning investigations, Conducting investigations, Processing data* and *Evaluating investigations*. Students' perceptions of their self-reflective and metacognitive skills were accessed through their comments about the worksheet on marking an investigation and they are briefly addressed. In the main, the analyses in this section are from pooled data and represent the views of the cohort, however, because Class SC performed the tasks in a different order from Classes TN and TC some analyses reflect this.

Third, the Chapter addresses assessment issues associated with the implementation of the different assessment regimes implemented in Classes TN, TC and SC. Students' perceptions of teacher assessed norm referenced, teacher assessed criterion referenced and student assessed criterion referenced assessment regimes are described. These analyses are at a class level because each class differed in their assessment regime.

Investigation Competencies

The Amount of Learning from Different Tasks

For each task, students were asked (Appendix M) about the amount of science content they had learned (Question 1) and the amount that they had learned about **doing** science investigations (Question 2). Sample questions are presented in Figure 14. Question 1 was intended to alert students to the difference between learning science content and learning about doing investigations. This distinction is also made by Hodson (1993, p. 106), when he said it is convenient to think about science education as having three major aspects; (a) learning science (b) learning about the nature and processes of science, and (c) learning how to do science. Students rated their learning from 0 to 5 with zero meaning 'nothing' and five meaning 'a lot' and the results are presented in Table 7.

Tasks	Questions
Worksheet	1. How much did you learn about writing hypotheses, working out the independent and dependent variables, how to control variables, etc?
	2. How much did you learn about **doing science investigations** from the worksheet?
Lipase investigation	1. How much did you learn about the reaction of lipase with milk?
	2. How much did you learn about **doing science investigations** from this activity?

Figure 14. Sample questions relating to the tasks

In terms of **learning science content knowledge**, the cohort gave the worksheet on terminology the highest average rating of 3.46 out of 5.00, followed by the catalyst investigation (3.25), the lipase investigation (3.08), the electromagnet

investigation (2.85) the teacher modelled investigation of an acid and calcium carbonate (2.82), the Panadol investigation (2.79), the ukulele investigation (2.74) and the worksheet on marking an investigation (2.69). In terms of **learning about doing science investigations**, the lipase investigation, with a mean rating of 3.15, was slightly ahead of the catalyst investigation (3.12) and the worksheet on terminology (3.12). These were followed by the ukulele investigation (3.04), the teacher modelled acid and calcium carbonate investigation (2.81) the Panadol investigation (2.58), the electromagnet investigation (2.46), the teacher modelled investigation of the pitch of a closed pipe (2.21) and the worksheet on marking an investigation (2.03).

Table 7 Students' mean rating of tasks in terms of how much content knowledge they learned and how much they learned about doing investigations (n = 66)

Tasks	Class	Learning science content knowledge: Class average	Learning science content knowledge: Cohort average	Learning about doing investigations: Class average	Learning about doing investigations: Cohort average
Worksheets					
Terminology	TN	3.77		3.68	
	TC	3.29		2.85	
	SC	3.32		2.82	
			3.46		3.12
Marking an investigation	TN	2.82		2.04	
	TC	2.55		1.89	
	SC	2.74		2.13	
			2.69		2.03
Teacher modelled investigations					
Acid and calcium carbonate	TN	3.00		2.82	
	TC	3.20		2.85	
	SC	2.59		2.77	
			2.82		2.81
Pitch of a closed pipe	TN	2.59		2.54	
	TC	1.86		2.00	
	SC	1.71		2.09	
			2.06		2.21
Student Investigations					
Lipase investigation	TN	3.38		3.67	
	TC	2.91		3.00	
	SC	2.95		2.82	
			3.08		3.15
Catalyst investigation	TN	3.86		3.81	
	TC	3.27		2.77	
	SC	2.62		2.81	
			3.25		3.12
Ukulele investigation	TN	3.14		3.18	
	TC	2.46		2.79	
	SC	2.67		3.33	
			2.74		3.04
Electromagnet investigation	TN	2.95		3.18	
	TC	3.08		2.47	
	SC	2.67		1.79	
			2.85		2.46
Panadol investigation	TN	3.09		2.91	
	TC	2.52		1.96	
	SC	3.33		2.94	
			2.79		2.58

Note 1 The tasks were rated on a 0 to 5 scale where 0 was 'nothing' and 5 'a lot'
2 Class TN and Class TC performed the worksheet on terminology, the lipase investigation, the teacher modelled acid and calcium carbonate investigation, the catalyst investigation, the teacher modelled pitch of a closed pipe investigation, the ukulele investigation, the worksheet on marking an investigation, the electromagnet investigation and the Panadol investigation, with the first four performed in May, the next four in August and the last task in November
3 Class SC performed the worksheet on terminology, the lipase investigation, the teacher modelled pitch of a closed pipe investigation, the ukulele investigation, the teacher modelled acid and calcium carbonate investigation, the catalyst investigation, the worksheet on marking an investigation, the Panadol investigation, and the electromagnet investigation with the first four performed in May, the next four in August and the last task in November

The following patterns are inferred from an analysis of data presented in Table 7. These data need to be interpreted with caution because students' mean ratings of the various tasks are detemined from their responses to one questionnaire item.

1. There was little difference between students' average rating of how much science content they learned and how much they learned about doing science investigations for each task. Students may have found this distinction difficult to make. The greatest difference was 0.66 for the worksheet on marking an investigation. This may have been due to the fact that the worksheet did not involve any laboratory work and therefore, students may have perceived that they learned little about doing investigations. The least difference between learning science content and learning to do science investigations was 0.01 for the teacher modelled acid and calcium carbonate investigation.

2. Students rated the teacher modelled investigation, factors affecting the pitch of a closed pipe, lowest in terms of learning science content knowledge and second lowest in contributing to learning about doing science investigations. Data presented later (Assertion 28. p. 113) suggests that students thought that they learned more when they performed investigations in a group.

3. Students rated their perceived learning of science content higher than their learning about doing science investigations for most tasks with the exception of the ukulele investigation, the lipase investigation and the teacher modelled pitch of a closed pipe investigation. For the ukulele investigation and the teacher modelled pitch of a closed pipe investigation students had to learn specific procedures to

collect the data. They used data logging software that was new to them, to analyse the frequency of a sound wave. This procedural knowledge may account for the high rating of these tasks. The lipase investigation may have rated highly because it was the first investigation in which the students participated.

4. Students indicated they learned more about doing investigations from the earlier investigations, the lipase investigation (3.15), the catalyst investigation (3.12) and the ukulele investigation (3.04) than from later investigations such as the electromagnet investigation (2.46) and the Panadol investigation (2.58). This is supported by qualitative data gathered in response to an open-ended question that asked students to give examples of their learning (Figure 15). These data illustrate students' perceptions that they learn more from new tasks and this notion is presented as Assertion 19. Examples of students' responses are presented to illustrate this finding.

Nikki's comment relates to the first investigation (Lipase investigation) and the other comments are about the last investigation (Panadol investigation). After five investigations it is evident that some students (Jo, Janet, Pippa and Olive) believed that they had gained little new knowledge from the Panadol investigation. The last comment by Anna, about this investigation, was atypical in that it indicated that she had learned most from the final investigation.

Nikki: Since it was the first investigation we learnt a lot because we didn't know how to do it & we hadn't done it before. (Lipase investigation)

Jo: I didn't learn much from doing this science investigation because I had learned all about doing science investigation in previous investigations. (Panadol investigation)

Janet: I already had a fair idea on what would happen in this investigation. I knew that the temperature would affect the rate it took to dissolve. I didn't learn much on how to do investigations because I'd already learnt my mistakes from my previous ones. (Panadol investigation)

Pippa: I think by now I have learnt what I need to know about science investigations in the classroom, and I didn't get much new knowledge out of this one. (Panadol investigation)

Olive: I didn't really learn much about doing science experiments because we had already done so much of these before, but I did learn that the hotter the temp of water the quicker the pannadol (sic) clears. (Panadol investigation)

Anna: I think that I learnt the most in this experiment although I didn't do it well. I didn't find that we had enough time at all. (Panadol investigation)

Assertion 19: Students perceive that they learn most from the first few investigations they perform.

Students' Perceptions of the Investigation Competencies Learned

For each task, students were asked to write examples of what they had learned about doing science investigations. Sample questions are shown in Figure 15. This section presents the categorisation of students' responses to these questions, tallies for each category, and examples of students' perceptions of what they learned about doing investigations. Some students gave quantitative responses such as "a lot" or "not much" instead of citing examples. These were not categorised because ratings of students' learning were addressed in the preceding section.

Task	Questions
Worksheet (Terminology)	Write down what you learned about **doing science investigations** from completing the worksheet and give some **examples.** Be honest. If you did not learn anything about doing science investigations then you should say why.
Lipase investigation	Write down what you learned about **doing science investigations** from the lipase investigation and give some **examples.** Be honest. If you did not learn anything about doing science investigations then you should say why.

Figure 15. A sample of the open-ended questions about the tasks students performed

The investigation competencies mentioned by students were similar to the assessment criteria that Class TC and Class SC used to evaluate their investigations (Appendix I). Therefore, this was used as a basis for categorising students' perceived competencies, however, some changes were made. These changes included (a) minor alterations to the wording of some categories and sub categories; (b) the addition of *Conducting investigations* sub categories, *Using equipment* and *Working safely* ; (c) the addition of a *Processing data* sub category, *Averaging results;* (d) the inclusion of new categories, *Writing reports* and *Managing time* , and (e) the deletion of the categories *Evaluating investigations* and *Creativity*.

It is acknowledged that some categories of competencies are not mutually exclusive. For example, being organised is linked with students' abilities to formulate an overall plan, to work cooperatively as a group to operationalise the plan, and to manage time effectively. Some comments were tallied in more than one category. For example, Pippa's response was tallied in the *Working cooperatively* category and the *Repeat trials* sub category. Kara's response was coded in the *Working cooperatively* category and the *Being accurate* sub category.

Pippa: ...to persevere. This was the first investigation in which our group worked well together. We did sufficient trials. (Ukulele investigation)

Kara: I knew what the outcome should be so that I learnt nothing in that respect but this experiment proved to me how you must work as a group and be organised. Also I found getting accurate results hard. (Catalyst investigation)

Table 8. Number of students who indicated that they had learned various investigation competencies for specific tasks (n = 66)

Investigation competencies	Tasks										
	Worksheets		Modelling		Student investigations						Total
	WT	WMI	TMI AC	TMI CP	LI	CI	UI	EI	PI		
Planning investigations											
writing hypotheses	26	2	0	3	2	0	1	2	1	37	
identifying variables	26	1	2	2	0	0	1	2	0	34	
overall planning	7	3	2	5	10	5	10	5	2	49	
preliminary trialing (refining plan)	5	2	2	2	5	2	0	0	3	21	141
Conducting investigations											
using equipment	0	0	3	11	2	1	10	0	1	28	
controlling variables	10	5	4	7	3	5	5	8	7	54	
repeating trials	2	0	6	5	7	3	5	0	7	35	
being accurate	2	1	12	1	13	10	3	6	5	53	
working safely	0	0	2	0	0	2	0	0	0	4	174
Processing data											
using tables	1	2	4	1	2	1	2	1	2	16	
averaging results	0	1	0	0	1	0	0	0	2	4	
using graphs	0	7	0	0	0	1	4	1	4	17	
drawing conclusions	1	2	0	0	1	0	0	0	0	4	41
Writing the report	0	6	0	5	1	2	3	4	2	-	23
Working cooperatively	0	0	0	0	2	6	0	0	2	-	10
Attending to detail	2	4	0	1	2	1	3	1	2	-	16
Managing time	0	2	0	0	2	4	0	4	6	-	18
Being organised	1	1	2	1	4	0	2	1	1	-	13

Note. The following abbreviations have been used

Worksheets

WT	Worksheet on terminology
WMI	Worksheet on marking investigations

Modelling

TMI AC	Teacher modelled investigation: Acid and calcium carbonate
TMI CP	Teacher modelled investigation: Pitch of a closed pipe

Investigations

LI	Lipase investigation
CI	Catalyst investigation
PI	Panadol investigation
UI	Ukulele investigation
EI	Electromagnet investigation

Table 8 indicates that there were 141 responses that mentioned *Planning investigations*, and *Overall planning* was mentioned more frequently than other sub competencies. It is not surprising that these competencies had a high frequency of response because this was a **new** aspect of experimental work that was being developed through the investigations. This is consistent with the finding previously presented; that perceived learning is greater from work completed early in the program (Assertion 19). In addition, students acknowledged that they learned about writing hypotheses, identifying variables, and preliminary trialing to improve experimental procedures. Twenty-six students identified the worksheet on terminology as contributing to their perceived learning about hypotheses and variables. This is expressed as Assertion 20. The lipase and ukulele investigations were mentioned by 10 students as contributing to developing overall planning skills.

Assertion 20: Students perceive that they learn most about writing hypotheses and identifying variables from the worksheet on terminology.

The following comments indicate typical perceptions of learning that were classified as ***Planning investigations***. Kit's statement about the worksheet on terminology was classified in the sub category *Identifying variables*. Her comments about the lipase and catalyst investigations were classified as learning about preliminary trialing and more specifically that preliminary trials can be used to refine experimental procedures. In addition, her comment about the catalyst investigation was classified as learning to perform repeated trials. In contrast, Kerry's statement about the lipase investigation revealed that she could see benefits in developing an overall plan before starting the investigation and this comment was classified in the *Overall planning* sub category. Statements by Jo, Gill and Olive are illustrative of students' perceived learning from the worksheet on terminology.

Kit: I learnt about the dependent & independent variables, as before I was confused about them & not sure which was which. (Worksheet on terminology).
I learnt that sometimes what you planned to do had to be changed when you actually did the experiment. (Lipase investigation)
I learnt about how useful preliminary trials are and that it is extremely important to control variables. About the importance of planning. The importance of doing the experiment more than once. (Catalyst investigation)

Kerry: ... and most importantly, that we had to know what we were doing Before (sic) starting the experiment (Worksheet on terminology).

Jo: I learnt how to write hypotheses, more correctly, how to control variables and what the independent and dependent variables are (Worksheet on terminology).

Gill: I already knew about hypotheses. I didn't know about variables until I completed this sheet (Worksheet on terminology).

Olive: I learnt a lot in this worksheet eg. all the meanings of words, planning, what to expect. It was also great guide so that whenever you got confused you could always refer back. I knew nothing about it before but now I feel totally confident with myself. This taught me a great deal (Worksheet on terminology).

The data (Table 8) indicate that students most frequently mentioned ***Conducting investigations*** (174 responses). Of these responses 54 related to *Controlling variables* and 53 to *Being accurate*. The following comments address the sub competencies *Using equipment* (Ann), *Controlling variables* (Ann and Kerry), *Repeating trials* (Anna), *Being accurate* (Anna and Penny) and *Working safely* (Penny).

Ann: I learnt how to use the computers. I learnt how to get the right sound wave also how to plan and correctly use equipment. I learnt how to cope with unusual results. I also learnt how to do a range graph. I also learnt how to keep my variables constant and also how to transport ideas into the experiment. (Ukulele investigation)

Kerry: ...we understood how to control variables, how to conduct the experiment properly, and how to watch for changes ... (Catalyst investigation)

Anna: I learnt that you have to be very accurate and precise when doing experiments. You have to keep all variables constant to eliminate any errors. Trials help to make results accurate and you always need to go over your experiment method and make sure there are no changes to be made before doing the actual experiment. (Catalyst investigation)

Penny: ... to take accurate measurements and to not mess about because things like hydrogen peroxide can burn. (Catalyst investigation)

Aspects of the investigations that related to ***Processing data*** were mentioned 41 times with 16 students commenting about *Using tables* and 17 about *Using graphs*. In the following comments Mary refers to learning to use tables correctly; Pearl to averaging results; Lara, Annette, Bron and Emma to developing graphing skills, and Annette and Kim to drawing conclusions.

Mary: ... to draw the table correctly and to write headings. (Panadol investigation)

Pearl: I learnt how to accurately collect data eg. take three samples and average them. (Lipase investigation)

Lara: I learnt how to do the graph. (Ukulele investigation)

Annette: ... not to be careless with titles on graphs. (Ukulele investigation)

Bron: I learnt that the graph must be the same as the table. (Ukulele investigation)

Emma: I learned that the best way to do this kind of graphing was in a line graph ... (Panadol investigation)

Annette: ... that to support a hypothesis I had to solve it, and... (Worksheet on marking an investigation)

Kim: Make sure you always have titles, a hypothesis and a conclusion. Make a conclusion say if the hypothesis was supported. (Worksheet on marking an investigation)

Interestingly, no students cited examples of learning about doing science investigations that could be classified as ***Evaluating investigations***. It was expected that students would be aware of the importance of this aspect of investigating from the *Evaluating findings* section on the Investigation Planning and Report Sheets (IPRS) (Appendixes D & F). It may be speculated that the lack of response was due to students interpreting **doing** to mean the activity part of the investigation and encompassing only planning and conducting. Later in the questionnaires, when students were asked to specifically document what they had learned about

Evaluating investigations their responses were characterised by a lack of detail. Therefore, the following assertion is made.

Assertion 21: Students do not consider *Evaluating investigations* as learning about investigation competencies.

Comments that students wrote about ***Writing the report*** for an investigation were superficial and are exemplified by Barbara, Henri and Anna.

Barbara: ... I learned about writing up investigations and performing practical work. (Ukulele investigation)

Henri: I learnt to plan and write up an experiment. (Ukulele investigation)

Anna: I already knew what the results would be but I learnt a lot about how to conduct an experiment and how to write it up. (Ukulele investigation)

Several students commented on the importance of ***Working cooperatively*** as a group.

Jan: ... that we had to all work together (my group). (Lipase investigation)

Gay: I learnt more about cooperation than how to conduct an experiment. (Panadol investigation)

Students mentioned ***Attending to detail*** in the experimental phase of the investigation and in the write-up. Comments by Annette and Terri are illustrative of this.

Annette: I learned from my mistakes that the tiniest detail can be important in the long run even if it doesn't seem important at the time. (Ukulele investigation)

Terri: We learned the scientific steps of experimentation. We learned how to explain (in detail) our experiment & also learned how to display our data visually. It taught me a very important lesson in how to write experiment details down properly & also taught me how to be as accurate as possible. (Ukulele investigation)

Students, for example Gill and Christie recognised the importance of ***Managing time*** to complete the investigation.

Gill: We learnt about planning out how to use the time given to us effectively. (Catalyst investigation)

Christie: To be more efficient with the investigation. (Electromagnet investigation)

Being organised seemed to be linked with the students' abilities to work cooperatively as a group and their abilities to operationalise their plan.

Annette: ... and that we really did have to be organised. (Ukulele investigation)

Harue: This was the best investigation. We were organised and worked quickly. I learnt how to conduct experiments successfully. (Panadol investigation)

The investigation competencies that students perceived they learned have been documented in this section and, in summary, the data are presented as Assertion 22. Many of these skills have been listed by Coles and Gott (1993). It is noticeable that students did not identify skills that associate one phase of an investigation with another, such as relating the conclusion to the hypothesis. These skills have been identified by numerous researchers (Bryce et al. 1991; Germann & Aram, 1996; Lawson, 1995) and, as such they form the basis for the justification of an holistic approach to investigating (Duggan et al. 1996; Hodson, 1992; Roth & Roychoudhury, 1993; Woolnough, 1989; Woolnough & Toh, 1990). The fact that students in this study did not identify relationships between the phases of an investigation is not surprising as researchers contend that higher order thinking is involved (Foulds et al. 1992; Kuhn et al. 1988; Lawson, 1995). It may be speculated that students are unlikely to do this unless prompted.

Another noticeable feature of the data is that students identified social and organisational skills associated with group work. This supports the claim by Linn and Burbules (1993) that group learning fosters social and workplace skills as well as cognitive skills.

Assertion 22: Investigation competencies that students claim to have learned include (a) planning investigations that comprises writing hypotheses, identifying variables, overall planning and preliminary trialing that can result in modifications to the plan (b) conducting investigations that comprises using equipment, controlling variables, repeating trials, being accurate and working safely (c) processing data that comprises using tables, averaging results, using graphs and drawing conclusions (d) writing reports (e) working cooperatively (f) attending to detail (g) managing time and (h) being organised.

Table 8 (p. 115) also reveals that different student investigations contributed differently to students' perceived learning of investigation competencies and this is presented as Assertion 23 (p. 124). The following discussion about the nature of the student investigations (the lipase, catalyst, ukulele, Panadol and electromagnet investigations) highlights different competencies that students thought they had developed. These data support the notions of task and domain specificity of learning, which have been identified as factors that need to be considered in performance/practical assessment (Gipps, 1994; Shavelson & Baxter, 1992).

For the **lipase investigation** students most frequently commented that they learned about *Being accurate* (13 students), *Overall planning* (10 students), *Repeating trials* (7 students), and *Being organised* (4 students) (Table 8, p. 115). The organisational skills required for this task were greater than had been required

for previous laboratory work. The students had experienced few quantitative laboratory exercises and the accuracy and care required to complete this task was new. Also, it was their first experience at planning an investigation and at conducting laboratory work over an extended period of time. The time taken for the lipase to break down milk fats into fatty acids (the dependent variable) was measured when an indicator signalled the formation of an acidic solution. Although students had used acid-base and starch indicators previously, the use of an indicator as a signal for something else, such as the time taken for a reaction was a new idea. On average, students perceived that they learned more about doing investigations from this task than from other tasks (Table 7, p. 110). However, students' comments need to be interpreted with caution because they may relate to weaknesses they identified in their performance, rather than real learning.

For the **catalyst investigation** students commented that they had learned about *Being accurate* (10 students), *Working cooperatively* (6 students), *Controlling variables* (5 students), *Overall planning* (5 students) and *Managing time* (4 students). The equipment in the catalyst investigation was difficult to set-up and manage, and students needed to take care to obtain accurate results. They were required to weigh-out the catalyst, transfer it to the reaction flask, measure the volume of the hydrogen peroxide, pour it in the reacting vessel and quickly replace the cork and start timing, and measure the volume of oxygen gas collected by the downward displacement of water. Cooperative group work was needed because one person added the catalyst to the reaction mixture and quickly replaced the stopper; the second held the water-filled inverted test tube to collect the oxygen gas, and the third timed the collection of the gas. Measuring the time for the reaction by the oxygen produced was not simple. Students either measured the volume of oxygen collected in a certain time, or measured the time taken to collect a certain volume of oxygen gas.

Students commented that their perceived learning from the **ukulele investigation** was about *Overall planning* (10 students), *Using equipment* (10 students), *Controlling variables* (5 students), *Repeating trials* (5 students) and *Using graphs* (4 students). The contribution of this investigation to learning about planning was likely to be high because it was early in the instructional program. *Using equipment* was mentioned by 10 students because, for the first time, they were required to use the computer to analyse the frequency of a sound wave. Graphing skills were mentioned by four students and this may be because the computer gave a frequency interval/range that students recorded and graphed as a bar. Students seemed to find this investigation easier than the others. There was no equipment to set up and also, once they were familiar with the computer program, the measurement of the dependent variable was simply a frequency interval read off the computer screen. From Gemma's comment it is evident that the way the dependent variable is measured influences the difficulty of an investigation.

Gemma: I learnt that its harder when you have to work out the way of measuring something than already having a machine and using it. (Ukulele investigation)

For the **Panadol investigation** students commented most about *Controlling variables* (7 students), *Repeating trials* (7 students), *Managing time* (6 students) and *Being accurate* (5 students). *Managing time* was mentioned more times than for other investigations probably because Panadol takes a considerable time to dissolve. Some students found it difficult to complete the investigation because they did not start their data collection early enough. In this investigation students needed to operationalise a definition of 'dissolving' because if they used low temperatures the last remaining 'bits' of Panadol did not completely dissolve. The investigation was quite straight forward and, as a result only two students commented that it had contributed to their learning about planning.

Students commented that the **electromagnet investigation** mainly contributed to their learning about *Controlling variables* (8 students), *Being accurate* (6 students) *Overall planning* (5 students) and *Writing the report* (4 students). Some students (Tammy and Jody) found this investigation difficult because they had to determine how best to measure the strength of the electromagnet. They had to operationalise their definition of magnetic strength. Most students used the distance that a paper staple moved and this provided an indirect measure of magnetic strength.

Tammy: Make sure what you test is measurable.

Jody: We need to find a more accurate way of measuring the distance of the staple from the electromagnet.

This discussion indicates that different investigations provide different learning opportunities for students. These qualitative data add insights to tallies of the number of students who reported that they learned particular competencies from different investigations, Table 8 (p. 115). Therefore, the notion of task and domain specificity of learning documented by other researchers (Gipps, 1994; Shavelson & Baxter, 1992) is supported by this research. Students' perceptions about the effect of the nature of the investigation on the competencies they learn is presented as Assertion 23.

Assertion 23: The nature of the investigation influences the investigation competencies that students believe they learn.

Aspects of the Cognitive Apprenticeship Model of Instruction

This section addresses aspects of the questionnaire that relate to students' perceptions of the cognitive apprenticeship model of instruction. It commences by ranking students' perceptions of the **best** ways to learn about doing investigations. This is followed by additional data about students' perceptions of articulating and coaching. Also, data are presented about teacher modelled investigations. Finally, there is a brief description of students' self-reflective and metacognitive skills as indicated by their responses to questions about the worksheet on marking an investigation.

The Best Way to Learn about Doing Investigations

In the November questionnaire students were asked to rank seven ways of learning in order of how they **best** learn to do investigations; doing an investigation in a group, talking with other students, watching and listening to the teacher model an investigation, talking with the teacher, doing an investigation by yourself, watching and listening to other student do an investigation, and correcting/marking an investigation. From this ranking, insights were gained about the relative contributions to learning about investigating made by (a) doing investigations, (b) articulating, (c) watching teacher modelled investigations, and (d) the assessment of investigations. Data about students' learning from teacher or student assessed investigations are discussed in the final section of this Chapter under the heading *Assessment.*

The questionnaire stated: Seven different **ways of learning** how to do investigations have been listed below. Which is the **best way for <u>you</u> to learn** how to do investigations? Rank the ways of learning from 1 (best) to 7 (worst). The average rankings are presented in Table 9.

Table 9. Mean cohort ranking of the way students best learn about doing investigations (n = 66)

Best way of learning	Mean ranking
Doing an investigation in a group	2.35
Talking with other students	3.52
Watching and listening to the teacher model an investigation	4.05
Talking with the teacher	4.12
Doing an investigation by yourself	4.43
Watching and listening to other students do an investigation	4.77
Correcting/marking an investigation	4.91

Note. The ways of learning were ranked from 1 (best) to 7 (worst).

Table 9 shows that, on average, students ranked *Doing an investigation in a group* as the best way to learn about investigating and this is expressed as Assertion 24.

Assertion 24: Students perceive that the best way to learn about doing investigations is by doing them with a group of students.

The perceived contribution of group work to learning is supported by this assertion and the ranking of *Doing an investigation by yourself* as fifth, Table 9. The rankings of *Talking with other students* and *Talking with the teacher* have implications for the teacher's role as a coach and are discussed in more detail in the *Articulating and coaching* section. The ranking of *Watching and listening to the teacher model an investigation* as sixth has implications for *Teacher modelled investigations* and is discussed in that section.

Articulating and Coaching

The ranking of *Talking with other students* (Table 9) indicates the perceived contribution that articulating with other group members makes to learning about doing investigations. Interestingly the ranking of learning from *Talking with the teacher* is not as high. Hence, students perceive that it is **better** to learn about doing investigations from talking with other students than from talking with the teacher. This ranking is supported by additional questionnaire data described below that access students' perceptions of the **helpfulness** of talking/articulating in learning about doing investigations.

Students responded quantitatively to the questions about the **helpfulness** of articulating (Figure 16) following the lessons during May and November. Students' ratings from zero (not at all) to five (a lot) were tallied and the means for each class and the cohort mean are presented in Table 10. Examples of students' perceived learning are also presented.

Questions
• How much did **talking with other students** in your group help you to learn about investigations?
• Give examples of what you learnt from **talking with other students** in your group.
• How much did **talking with the teacher** help you to learn?
• Give examples of what you learnt from **talking with the teacher.**

Figure 16. Questions during May and November about the helpfulness of learning from articulating

Table 10. Mean rating of the amount of learning about investigating from talking with other students and from talking with the teacher for the Classes TN (n = 20), TC (n = 21) and SC (n = 21), and for the cohort (n = 62) during May and November

		Talking with other students		Talking with the teacher	
Time	Class	Class mean	Cohort mean	Class mean	Cohort mean
May	TN	3.86		3.19	
	TC	3.59		3.00	
	SC	3.32		2.82	
			3.58		3.00
November	TN	2.90		2.78	
	TC	3.22		1.61	
	SC	3.08		2.40	
			3.07		2.25

Note. The students rated ways of learning on 0-5 scale where 0 was 'nothing' and 5 was 'a lot'

An emerging trend (Table 10) during May and November, was that on average, students perceived that they learn more from talking with students than from talking with the teacher. *Talking with other students* and *Talking with the teacher* were respectively 3.58 and 3.00 during May, and respectively 3.07 and 2.25 during November. Most students were positive about their learning from talking with other students. Some, for example Kara and Barbara, said that there were benefits in having different points of view because the input was greater. Lib, Gay and Cara's comments indicate that the articulation of ideas helps them to learn. Emma acknowledges that fellow students explain more clearly and are able to devote more time to explanations. The comment by Amy shows that the consensus of others helps her to feel confident.

Kara: Talking in my group helped a lot with understanding the investigation as you get more than one point of view.

Barbara: Making your options greater. They think of things you may not. The input is greater.

Lib: They helped me understand what we were doing, they helped me set up the apparatus.

Gay: It's easier to do things with your group. You understand things a lot more.

Cara: I learned that talking to another student makes it easier to do the experiments because we can talk about our mistakes.

Emma: They could explain things more clearly and understand, and conserntrate (sic) more on you than the teacher with many other questions to think about.

Amy: Talking with other students in your group gives you confidence in what you write and helps you to cooperate in organising an investigation. They are also good to talk with about understanding results. It helps you learn off each other.

However a minority of students (Henri, May, Kath, and Jess) found working in a group difficult because of the group dynamics. The notion that not all students like group learning is consistent with research by Linn and Burbules (1993) and Roychoudhury and Roth (1996). Linn and Burbules report that group learning may be unproductive for "learners who have dysfunctional views of group interaction" (p. 114), and Roychoudhury and Roth report that two students out of 46 junior high school physics students preferred to work alone. Roychoudhury and Roth added that most students "acknowledged the benefits of 'pooling' ideas and effort". This contrasts with the views of Pippa as she stated that pooled ignorance does not always contribute to learning. Linn and Burbules address this perspective and state that learning outcomes will not be achievable if students are unable to access appropriate information.

Henri: I learnt that it is often hard to agree on matters.

May: It helped a bit but our group didn't work very well together so our communication wasn't great.

Kath: One member of our group left us because she reconed (sic) what we were doing was wrong.

Jess: That we all have different ideas about how we think we can do it and what is right. Talking doesn't always solve problems.

Pippa: We didn't learn much from each other because we knew all the same things.

Although some students indicated that not all groups functioned well and that student-student interactions are not always fruitful, Assertion 25 is supported by the following two interpretations of the questionnaire data: Students perceive that it is **better** to learn about doing an investigation from talking with students than from talking with the teacher (Table 9); and students perceive that the **amount** of learning from talking with students is greater than from talking with the teacher (Table 10). As well, students' perceptions that the best way to do investigations is in a group (Table 9, Assertion 24) may partly be attributable to the contribution of student-student interactions to learning.

Assertion 25: Most students perceive that they learn more about doing investigations from talking with their peers than from talking with their teacher.

This assertion does not indicate the sorts of learning that occur through student-student interactions. Previous studies (Christensen & McRobbie, 1994) indicate that the questions students asked teachers were mainly focused on the procedures to complete the task. Assertion 25 reflects poorly on the teacher's role as 'coach' during science investigations and raises doubts about the teachers' abilities to implement informal, or unstructured formative assessment as described by Harris and Bell (1994), and Radnor and Shaw (1995). It may be that teachers are too busy with classroom management issues to act as effective coaches or that they are ineffective at gathering appropriate feedback from students from which to make judgements about the nature and type of help they should give students.

A second trend to emerge from Table 10 (p. 128), regardless of whether students' talk was with their peers or teacher, was that their perceived amount of learning from discourse was greater during May than November and this is presented as Assertion 26. During May, *Talking with other students* and *Talking with the teacher* were respectively 3.58 and 3.00, and during November they were respectively 3.07 and 2.25. This supports Assertion 19, that students perceive they learn most from the first investigations they performed.

Assertion 26: Most students perceive that they learn more about investigations from talking with their peers and their teacher early in the instructional program.

For Class TC the decrease in students' learning (Table 10) from **talking with the teacher** from 3.00 (May) to 1.61 (November) is more than for students in other classes. It appears that Miss Mills thought the students were capable of conducting the last investigation, the Panadol investigation, by themselves and encouraged students to solve their own problems. Many students in this class had the same view.

Sas: I didn't learn anything from the teacher because the whole point of an investigation is to work something out for yourself. (November)

Olive: The teacher didn't help us with this as she wanted us to work independently. (November)

To some extent students' responses to the instruction, "Give examples of what you learnt from talking with the teacher," revealed their preferences for particular teaching characteristics and their expectations of teachers' roles (Assertion 27). It seemed that most students expected direct answers to questions they asked the teacher; that is, explicit responses from the teacher directing them what to do or telling them the answer. When teachers helped students to think about and find their own solutions to difficulties, instead of merely providing answers,

some students expressed dissatisfaction. The contrasting comments about Miss Mills and Mr Brogo by Kerry, show her preference for Miss Mills' full explanations rather than Mr Brogo's attempts to get the students to work out their own solutions. On the other hand, Claire states that she prefers Mr Brogo's approach because he makes her work things out for herself. Like Kerry, Phyllis prefers Miss Mills' explanations. Clearly Phyllis prefers considerable structure and scaffolding in the teacher's responses to her questions.

Kerry: The teacher (Miss Mills) explained things in a clearer way, I learnt more in general. It's hard to say something specific.

Kerry: Didn't talk to the teacher because he (Mr Brogo) wouldn't give us a suitable answer.

Claire: The teacher (Mr Brogo) helped you but made you work it out for yourself which was good.

Phyllis The teacher (Miss Mills) showed us how to prepare for the experiment well. She took us through the steps.

Assertion 27: Some students prefer teachers who give full explanations and some prefer teachers who encourage and assist them to solve their own problems.

Students gave varied examples of their learning from talking with the teacher. Some students, for example Pam, stated specific details; others such as Kara and Terri said that the teacher gave them a different perspective of the investigation; and others still said that they didn't talk to the teacher at all.

Pam: How to write up experiments well. How to plan experiments etc.

Kara: Talking with the teacher helped us a lot to understand the investigations and we could see errors we made and didn't realise.

Terri: It gave us different views of the investigation also giving us different variables to look at.

Teacher Modelled Investigations

Teacher modelling compared with student investigations

Table 9 (p. 126) indicates that students ranked *Watching and listening to the teacher model an investigation* behind *Doing an investigation in a group* and *Talking with other students* in terms of the **best** way to learn about doing investigations. Table 7 (p. 110) indicates that students rated the teacher modelled pitch of a closed pipe investigation last (eighth) and the teacher modelled acid and calcium carbonate investigation fifth in terms of **how much** the tasks contributed to learning about doing science investigations. The following quotations support the quantitative data and indicate that, as a result of observing teacher modelled investigations, students perceived that not a great deal of learning occurred. These ideas are expressed in Assertion 28.

Kim: I learned not much. It was boring, because we just watched. I didn't know what was happening and couldn't see. (Pitch of a closed pipe)

Kath: I didn't learn much because the teacher did all the work and didn't envolve (sic) the students much. (Acid and calcium carbonate)

Cara: I didn't learn very much from this one, maybe because the teacher did it and not the students. I also learned that the higher the consentration (sic) of acid the quicker the reaction takes place. (Acid and calcium carbonate)

Sara: I learned quite a lot about doing science investigations, although I think I would have learned more if I had carried out the investigation. (Acid and calcium carbonate)

Angela: I didn't learn as much as I think I would have if I had done it myself - instead of watching the teacher (Acid and calcium carbonate)

Assertion 28: Students perceive that they learn less about doing investigations from teacher modelled investigations than from student investigations.

Teacher modelling compared with student investigations and worksheets for learning about the phases of investigating

Data were gathered to determine the contribution of teacher modelled investigations to learning about the phases of investigating; *Planning investigations, Conducting investigations, Processing data,* and *Evaluating investigations.* The data were collected at the end of the instruction during May and August. It indicated from which task students perceived they had learned most about the phases of investigating; a worksheet, a teacher modelled investigation or two student investigations. Students' responses were tallied and the percentage of students who indicated that they had learned most from a particular task are presented in Tables 11, 12, 13, and 14. The May percentages are presented in Table 11 (Classes TN and TC) and Table 12 (Class SC). The August percentages are presented in Table 13 (Classes TN and TC) and Table 14 (Class SC). Different tables are used because Class SC did the tasks in a different sequence.

Table 11. Percentages of students in Class TN and Class TC who nominated particular tasks as most effective in helping them to learn about Planning investigations, Conducting investigations, Processing data and Evaluating investigations during May (n = 40)

Task	Percentage of students			
	Planning investigations	Conducting investigations	Processing data	Evaluating investigations
Worksheet				
Terminology	10	5	8	3
Modelled investigation				
Acid and carbonate	12	7	13	15
Student Investigations				
Lipase investigation	21	21	28	15
Catalyst investigation	57	67	51	67

Table 12. Percentages of students in Class SC who nominated particular tasks as most effective in helping them to learn about Planning investigations, Conducting investigations, Processing data and Evaluating investigations during May (n = 19)

Task	Percentage of students Planning investigations	Conducting investigations	Processing data	Evaluating investigations
Worksheet				
Terminology	5	0	7	8
Modelled investigation				
Pitch of closed pipe	5	0	7	8
Student Investigations				
Lipase investigation	16	16	7	15
Ukulele investigation	74	84	79	69

Table 13. Percentages of students in Class TN and Class TC who nominated particular tasks as most effective in helping them to learn about Planning investigations, Conducting investigations, Processing data and Evaluating investigations during August (n = 45)

Task	Percentage of students Planning investigations	Conducting investigations	Processing data	Evaluating investigations
Worksheet				
Marking an Investigation	2	4	9	5
Modelled investigation				
Pitch of closed pipe	2	4	5	5
Student Investigations				
Ukulele investigation	31	57	40	44
Electromagnet investigation	64	35	47	46

Table 14. Percentages of students in Class SC who nominated particular tasks as most effective in helping them to learn about Planning investigations, Conducting investigations, Processing data and Evaluating investigations during August (n = 23)

Task	Percentage of students			
	Planning investigations	Conducting investigations	Processing data	Evaluating investigations
Worksheet				
Marking an Investigation	9	14	18	5
Modelled investigation				
Acid and carbonate	17	9	9	9
Student Investigations				
Catalyst investigation	22	41	55	41
Panadol investigation	52	36	18	45

From the tables it is evident that students perceived that the teacher modelled investigations contributed less to their learning about *Planning investigations*, *Conducting investigations*, *Processing data* and *Evaluating investigations* than the student investigations (Assertion 29). These data are consistent with previously presented data that indicated that students learned less about investigations as a whole from teacher modelled investigations than from investigations they performed themselves (Assertion 28). The data also imply that worksheets are poor at developing investigation competencies, however, this may be because they focused on a narrow range of skills.

Assertion 29: Students perceive that teacher modelled investigations contribute less than student investigations to their learning about *Planning investigations*, *Conducting investigations*, *Processing data* and *Evaluating investigations*.

Examples of learning from teacher modelling

Although Table 7 (p. 110) indicates that from teacher modelled investigations the amount of learning was low, when students were asked to give specific examples of their learning from teacher modelled investigations (Table 8, p. 115) their examples focused more on *Conducting investigations* than on other competencies. For the closed pipe investigation students most commonly cited the sub competency *Using equipment* (11 students). Students' references to using equipment were significant because, for the first time, they used the computer to analyse the frequency of a sound wave. For the teacher modelled acid and carbonate investigation students perceived that they learned about *Being accurate* (12 students) and this competency was an important aspect of the data collection. The teacher had to accurately control variables such as temperature, volume of acid and mass of calcium carbonate, as well as measure the time taken for limewater to go milky. For many students this was one of their first experiences in quantitative chemistry and this is likely to have contributed to their questionnaire response. The notion that students learn most about the investigation competency *Conducting investigations* from the teacher modelled investigations and that the emphasis on the sub competencies varies with the context of the investigation is expressed as Assertion 30.

Assertion 30: Students perceive that for teacher modelled investigations the competency they learn most about is *Conducting investigations*, and for different investigations they learn different sub competencies associated with *Conducting investigations*.

These data raise concerns about the teacher modelling aspect of the cognitive apprenticeship model of instruction as implemented in this study. Modelling is a crucial element of the instructional approach as it provides the basis for the conceptual model of the target process (Collins et al. 1989). Also, according to Lave, in Collins et al. (1989) it provides an advanced organiser for the process, and a

guide for independent practice. In this study it may be that teacher modelling was important in providing a conceptual model and in guiding individual student practice even though students did not recognise its contribution to learning.

Self-Reflective and Metacognitive Skills

Tasks that had the potential to develop students' self-reflective and metacognitive skills included the questionnaire, the section of the IPRS that asked students to evaluate their investigation, the worksheet on marking an investigation, and for Class SC, the self-assessment of their science investigation. Students were not asked about the effect of the questionnaire on their learning. With regard to the evaluation section of the IPRS it has been reported previously that students do not consider *Evaluating investigations* as learning about investigation competencies, Assertion 21. Class SC's perceptions of student assessed criterion referenced assessment are discussed in the next section of the Chapter. The following discussion addresses students' perceptions of the worksheet on marking an investigation.

The purpose of the worksheet on marking an investigation was to encourage students to reflect on an investigation performed and written up by other students, and as a consequence, to develop a more reflective approach to their own work. The worksheet required students to assess an investigation (factors affecting heart beat) completed by previous Year 9 students. Major weaknesses in the write-up of the investigation included the limited control of variables, the failure to average results and the subsequent use of graphs for each data set, and the lack of detail in the written report. Table 8 (p. 115) indicates that from this task students learned about *Controlling variables* (5 students), *Using graphs* (7 students), *Writing the report* (6 students) and *Attending to detail* (4 students). Pippa, Barbara and Annette identified graphing as a major weakness in the investigation. Statements by Lib,

Pippa and Barbara revealed an increased awareness of the difficulties associated with assessing work. In contrast, many students could see no point in the task and this is typified by Gay's comment.

Pippa: I found that the worksheet exercise didn't teach me much about science investigations, just about how difficult it is to mark them. The only thing I learnt was to label everything and to graph averages, not every single result.

Barbara: In marking this sheet I learnt it must be hard for a teacher to mark disorganised work. Therefore I now know: Graphs should show averages not specific results, you must be able to control (test) your hypothesis, one must title all graphs and tables, all variables must stay the same except the independent.

Annette: With the investigation that we marked I found out a lot of things that the people did wrong I oftend (sic) did wrong myself. So I learnt that to support a hypothesis I had to solve it, and only to graph averages. I also learnt from their mistakes.

Lib: I understand what teachers are looking for in our investigations but generally I thought this sheet was boring.

Gay: I didn't see why this was important for me. (Sorry if it sounds rude!) I didn't go through it properly and therefore didn't benefit from this.

Quantitative data presented in Tables 13 and 14 compare students' perceptions of the worksheet on marking an investigation, the teacher modelled investigations and the student investigations, in terms of effectiveness in helping them to learn about the phases of investigating. On average, students in all classes identified the worksheet as more effective in helping them to learn about *Processing data* than the teacher modelled investigations, and Class SC also indicated that the worksheet was more helpful in their learning about *Conducting investigations*. Students, however, indicated that the effectiveness of learning from the student investigations was considerably higher than the other tasks across the four phases of investigating.

It may be that students are insufficiently aware of how they learn to recognise the benefits from self-reflective and metacognitive tasks. It is likely that they need explicit instruction in these thinking skills to be able to appreciate possible benefits from them.

Assessment

For the third questionnaire, depending on their class, students were asked to compare their learning from either "correcting an investigation" or "marking an investigation" with other ways of learning to do investigations. The other ways of learning about doing investigations included the following; doing an investigation in a group, talking with other students, watching and listening to the teacher model an investigation, talking with the teacher, doing an investigation by yourself, and watching and listening to other students do an investigation. Data were pooled across classes in Table 9 (p. 126), to indicate that when asked to rank these ways of learning in order from "the **best** way for you to learn how to do investigations" to the worst, on average, students across classes ranked correcting and marking an investigation last. This ranking does not necessarily indicate the amount of learning from the assessment of their work. The notion that students perceive that the worst way to learn about doing investigations is from correcting or marking an investigation is expressed as Assertion 31. Perhaps this is because they perceived that the investigations had been completed and that making corrections is a waste of time and/or that it serves no purpose.

Assertion 31: Students perceive that the worst way to learn about doing investigations is by correcting or marking investigations.

At the class level students were asked to rank different aspects of learning from assessment. This was because, as described previously in Chapter 3, different assessment procedures were used for the three classes. The investigations for students in Classes TN and TC were teacher assessed and students were asked to rank "correcting an investigation" because they were given this opportunity following the teacher assessment of their investigation. The students in Class SC assessed their own investigations by matching them to master answer sheets

prepared by the researcher, and assigning themselves grades for stated criteria. They were asked to rank "marking an investigation".

The individual class rankings are presented in Table 15. Class TN's mean ranking was highest and Class SC's was lowest. Reasons for these rankings were not explored in this study. The following suggestions for the low ranking of Class SC are proposed with caution. Students in Class SC may have found self-assessment difficult, or boring and/or they learned little, and consequently they ranked it as the worst way to learn about doing investigations. Also, they may have perceived that their teachers lacked interest in their progress because they had not assessed their work and hence, they responded by a lack of interest in improving their performance.

Table 15. Class assessment mean ranking of the best way to learn about doing investigations

Class	Assessment	Mean ranking
TN (n=21)	Correcting an investigation	3.90
TC (n= 22)	Correcting an investigation	5.19
SC (n= 23)	Marking an investigation	5.95

Note. Ways of learning from correcting/marking investigations ranked from 1 (best) to 7 (worst).

Additional questions to obtain information about how students responded to the assessment procedures are presented in Figure 17 and data were gathered after the lessons during May and November. Students indicated **how much** they learned from their assessment regime from zero (not at all) to five (a lot). The ratings were tallied and the class means are presented in Table 16.

Class	Questions
TN	How much did the **teacher's marking** of the investigations help you to learn? Give examples of what you learnt from the **teacher's marking** of the investigations.
TC	How much did the **teacher's marking** of the investigations help you to learn? Give examples of what you learnt from the **teacher's marking** of the investigations.
SC	How much did **marking your own investigations** help you to learn? Give examples of what you learnt from assessing your own investigations.

Figure 17. Questions asked about the assessment of the Investigation Planning and Report Sheets

Table 16. Mean rating of how much was learned from the different assessment procedures and the student investigations for Class TN, Class TC and Class SC during May and November

Class	May		November	
	Assessment	Student investigations	Assessment	Student investigations
TN	2.60	3.74	3.00	2.91
TC	2.55	2.89	2.22	1.96
SC	2.15	3.08	1.73	1.79

Note. Learning from the assessment rated from 0 (nothing) to 5 (a lot).

The data (Table 16) indicating **how much** was learned from the assessment are consistent with data that ranks the **best** way for students to learn (Table 9 & Table 15). In comparison with the other classes, Class TN (Table 16) **rated** their assessment procedure during May (2.60) and November (3.00) as contributing **more**

to learning about doing investigations and **ranked** it (3.90, Table 15) as a **better** way to learn than did Classes TC and SC. Class SC **rated** their assessment procedure during May (2.15) and November (1.73) as contributing least to their learning (Table 16) and **ranked** it worst in terms of ways to learn about doing investigations (5.97, Table 15). These data indicate that students from Class SC perceive self-assessment contributes less to their learning than students in Classes TN and TC who had their investigations assessed by a teacher. This is presented as Assertion 32.

Assertion 32: Students who assess their own investigations perceive that they learn less from this process than students who have their investigations assessed by teachers.

In addition, Table 16 provides a comparison of students' ratings of **how much** they learned about doing investigations from the assessment and from the student investigations they performed during May and November (Table 16). During May students rated the assessment as contributing less to their learning than student investigations. During November, Class SC indicated that they learned marginally less from the assessment than from the student investigations, however, Classes TN and TC indicated that they had learned slightly more from the assessment. The May assessment data are consistent with previous data that revealed students' low opinions of learning from assessment. The November data are difficult to explain. They may merely indicate that students were disinterested in performing the November investigations because they were at the end of the program and because the end of the school year was approaching. This is supported by earlier data (Assertion 19) which shows that students perceive that they learn more from the first few investigations they perform.

The following discussion is based on examples of learning that students provided in response to the questions about assessment, Figure 17. Students in Classes TN and TC were asked to give examples of what they learned from the teacher's marking of their investigation and students in Class SC were asked to give examples of what they had learned from their own assessment.

The comments made by students in **Class TN** (teacher assessed norm referenced assessment) related to the (a) lack of learning from the assessment procedure, (b) the helpfulness of whole-class feedback when students received their assessed work and (c) the fact that they were able to see where they had made mistakes.

(a) lack of learning

Dolly: I didn't learn much from her marking. (May)
I didn't really find that the teacher's marking affects me, accept (sic) that I need to try harder next time. (November)

May: I didn't learn much as there was no teacher comment on my mark.

Jo: Didn't learn much, except make sure you fill in <u>all</u> of the questions.

(b) whole-class feedback

Anna: She did not really write down much in our sheets but she explained a lot in class.

May: It helped a bit but it was more helpful when the teacher went through it with the class as a whole and told us what to include.

(c) identifying mistakes

Annette: I learned that they (teachers) mark your problems and help you see what you did wrong.

Anne: We were able to see the mistakes that we had missed when we were reviewing it. We then are able to know what to do the next time.

Cara: I was able to see where I went wrong.

Some students in **Class TC** (teacher assessed criterion referenced assessment) made comments about the lack of learning from the assessment and others said that it was beneficial in helping them to identify mistakes.

(a) lack of learning

Claire: It didn't really help much. (May)

Penny: Not very much because I couldn't really understand what she meant. (May)

Susan: Nothing. (November)

(b) identifying mistakes

Kit: That when you draw tables you only put the units in the heading squares. That I should control variables better. (May)

Em: I learnt a lot from them marking it because I was reading through there (sic) comments I understood what I had done wrong. (May)

Kerry: We understood where we went wrong and what we could do to fix it up. (May)

Susan: What not to do again. (May)

Claire: To expand my answers longer. (November)

Kerry: Be precise - what I did right and wrong. (November)

Kit: It was good to see what we had done wrong and what to improve on next time. (November)

Em: They wrote how we could make it better & explained it much more clear. (November)

Penny: You learnt where you went wrong and why you got the mark you did. (November)

Simone: It gave you an idea on what the teacher is looking for you write about the investigation. (November)

In **Class SC** (student assessed criterion referenced assessment) some students such as Alice said that the assessment process made no difference and others, for example Helen, stated that it wasn't necessary because they had already realised what they had done wrong. Belinda, Rose and Beth were more positive saying that

they could identify mistakes. A few students, Gill, Pippa and Clara recognised the self-reflective component of the self-assessment.

(a) lack of learning

Alice: It didn't make any difference to me. (May)

Helen: It wasn't necessary as we already knew what we had done wrong and what we had done right. (May)

(b) identifying mistakes

Belinda: Find out what your mistakes were. (May)

Rose: Learnt from my mistakes. (November)

Beth: To learn from your mistakes because your (sic) going through the answer and marking it yourself. (November)

(c) critical analysis

Gill: To be critical of our own work. (November)

Pippa: It was good to criticise yourself so you could improve in that area next time. (November)

Clara: Marking your work gives you more understanding of what your teacher wants. (November)

In summary, two assertions can be developed from the examples of learning that students provided in response to questions about what they had learned from the assessment, Figure 17. First, the students in Class TC and Class SC did not mention using the criteria to assign grades. This was the students first science experience with this form of assessment so it was expected that some comment would be made about the procedure. Students' failure to acknowledge this aspect of assessment is expressed as Assertion 33. The espoused strengths of criterion referenced assessment are numerous (Biggs & Moore, 1993; Gipps, 1994; Harris & Bell, 1994; Popham, 1992) so it was expected that some reference to the assessment would be made.

Assertion 33: Students do not acknowledge that criterion referenced aspects of assessment are helpful in providing feedback on investigation performance.

Second, for each of the classes there was a wide range of viewpoints about students' perceived learning from the assessment of their investigations. This idea is presented as Assertion 34.

Assertion 34: Students' perceptions of the amount of learning resulting from teacher and student assessment vary widely.

For the final questionnaire during November, students were asked whether they preferred teacher marking or student marking of their investigations. From Table 17 it is noted that almost half the students in Class SC said that they preferred student marking (Assertion 35). This is interesting in view of the previous finding that Class SC learned least from the assessment of their work (Assertion 32). The reasons for this are not clear, however, the low preference for student marking from Class TN and Class TC does indicate that if students are not exposed to a particular form of assessment then they are unlikely to recognise its value.

Assertion 35: Students who have not assessed their own work are less likely to see the value in this assessment procedure than students who have.

Table 17. Percentages of students in Class TN, Class TC and Class SC who indicated a preference for either teacher assessment or student assessment

	Percentages of students	
Class	Teacher assessment	Student assessment
TN (n = 21)	90.5	9.5
TC (n = 22)	86.4	13.6
SC (n = 23)	52.2	47.8

Summary of the Chapter

This Chapter addressed the results of three questionnaires that students completed after instruction during May, August and November. Both quantitative and qualitative data have been presented and 16 assertions have been formulated. The assertions (Figure 18) have been clustered according to investigation competencies, aspects of the cognitive apprenticeship model of instruction, and the assessment regimes, although it is acknowledged that the clusters are not necessarily mutually exclusive. This clustering provides a framework on which to report research findings and to address the research questions.

In terms of investigation competency, students identified isolated skills that they had learned but did not indicate that they had acquired competencies that relate one phase of an investigation to another, such as relating their conclusion to the hypothesis. In addition, they identified social and workplace skills (Assertion 22). Students' comments reinforced the notion of task and domain specificity and this has implications for performance assessment as different tasks test different skills or competencies (Assertion 23). Learning most from the first few investigations indicates the impact of novel tasks on students' learning (Assertion 19).

The data presented about the cognitive apprenticeship model of instruction raised some concerns. These concerns need to be considered in the light of the pretest and posttest data that showed strong gains in students' investigation competencies. Students perceived that they learned more from talking with their peers than their teacher (Assertion 26). This raises concerns about the teachers' role as a coach and the impact of formative assessment. Similarly the impact of teacher modelled investigations is questioned as students indicated that they learned less from teacher modelled investigations than from doing the investigations (Assertion 28).

With regard to assessment, students perceived that the worst way to learn about doing investigations was by correcting or marking an investigation (Assertion 3[illegible]). Students who assessed their own work, however, perceived that they had learned less from this process than students who had their investigations assessed by their teacher (Assertion 32). Furthermore, students did not acknowledge that criterion referenced assessment was of benefit to them (Assertion 33).

Assertions

Investigation competencies

19 Students perceive that they learn most from the first few investigations they perform

20 Students perceive that they learn most about writing hypotheses and identifying variables from the worksheet on terminology

21 Students do not consider *Evaluating investigations* as learning about investigation competencies

22 Investigation competencies that students claim to have learned include (a) planning investigations that comprises writing hypotheses, identifying variables, overall planning and preliminary trialing that can result in modifications to the plan (b) conducting investigations that comprises using equipment, controlling variables, repeating trials, being accurate and working safely (c) processing data that comprises using tables, averaging results, using graphs and drawing conclusions (d) writing reports (e) working cooperatively (f) attending to detail (g) managing time and (h) being organised

23 The nature of the investigation influences the investigation competencies that students believe they learn.

Aspects of the cognitive apprenticeship model of instruction

(a) Articulating and coaching

24 Students perceive that the best way to learn about doing investigations is by doing them with a group of students.

25 Most students perceive that they learn more about doing investigations from talking with their peers than from talking with their teacher

26 Most students perceive that they learn more about doing investigations from talking with their peers and their teacher early in the instructional program

27 Some students prefer teachers who give full explanations and some prefer teachers who encourage and assist them to solve their own problems

(b) Teacher modelled investigations

28 Students perceive that they learn less about doing investigations from teacher modelled investigations than from student investigations

29 Students perceive that teacher modelled investigations contribute less than student investigations to their learning about *Planning investigations*, *Conducting investigations*, *Processing data* and *Evaluating investigations*.

30 Students perceive that for the teacher modelled investigations the investigation competency they learn most about is *Conducting investigations*, and for different investigations they learn different sub competencies associated with *Conducting investigations*

Assessment

31 Students perceive that the worst way to learn about doing investigations is by correcting or marking investigations.

32 Students who assess their own investigations perceive that they learn less from this process than students who have their investigations assessed by teachers

33 Students do not acknowledge that criterion referenced aspects of assessment are helpful in providing feedback on investigation performance.

34 Students' perceptions of the amount of learning resulting from teacher and student assessment vary widely.

35 Students who have not assessed their own work are less likely to see the value in this assessment procedure than students who have.

Figure 18. Summary of assertions from student questionnaires

CHAPTER 6

TEACHER INTERVIEWS

Overview of the Chapter

This Chapter discusses data gathered from semi-structured interviews with the teachers who participated in the study. The teacher interview data relates to the three themes associated with the research questions; the development of investigation competencies, aspects of the cognitive apprenticeship model of instruction, and different assessment regimes. Investigation competencies are considered in terms of teachers' perceptions of the pretest (lipase investigation) and of the development of students' investigation competencies throughout the program. In the section on the cognitive apprenticeship model of instruction, teacher modelled investigations are considered along with teacher guidance, scaffolding and fading, and self-reflective and metacognitive skills. Finally, teachers' perceptions of the assessment regimes; teacher assessed norm referenced, teacher assessed criterion referenced, and student assessed criterion referenced assessment are considered.

The teachers were interviewed individually on two occasions during the May investigations; after they had modelled an investigation and at the completion of the 10 lesson sequence. For the second and third sequence of 10 lessons during August and November, teachers were interviewed at the completion of the lesson sequences. During the interviews, the researcher made hand written scripts of the teachers' comments. The scripts were typed into notes and shown to the teachers to ensure they represented an accurate record. No changes to the scripts resulted from this procedure. The teachers rotated amongst classes after the 10 lesson segments (Appendix B) so that the students had subject specialist teachers for chemistry, physics and biology units of work.

Investigation Competencies

Teachers' Perceptions of Students' Investigation Competencies from the Pretest

Teachers' responses to the pretest (lipase) investigation support some of the data gathered from the pretest Investigation Planning and Report Sheets (Chapter 4). The teachers' comments about the pretest provide a snapshot of the students' first attempts at investigations of the type described in this research (Appendixes D & E).

Teachers found the pretest lessons difficult because they were asked to not help the students. "Letting them have free reign was difficult knowing that they were making so many mistakes," said Miss Mills and she added, "This is not the way I would normally teach." Mrs Grant said that she "...perceived the students' discomfort," because she believed that they felt lost and hopeless at dealing with the task. Mrs Cross added that "It was hard to know what to tell them." The teachers were surprised and disappointed that their students faired poorly on the pretest investigation, however, poor performances should have been anticipated because this was the students' first encounter with an open investigation. In addition, they were uncomfortable with their role as an observer, rather than a helper or guide.

In terms of *Planning investigations*, teachers said that the students were very disorganised and that their plans lacked detail and coherence. "They were unable to outline the procedural steps," said Mrs Grant and she added, "Their plans were very badly written and you couldn't use them as a recipe. They did not specify amounts of substances that they would use, how they planned to measure the variables, or the variables that they would control." These ideas are expressed later in Assertion 37. These comments suggest that the students lacked an overall plan and did not attend to necessary details. Also, they did not plan for the collection of quantitative data, a characteristic previously reported by Black (1990). The lack of planning associated

with how to measure the variables is consistent with a study by Duggan et al. (1996) in which they report difficulties that 12-14 year old students experience operationalising continuous independent variables.

With regard to *Conducting investigations*, Miss Mills highlighted that students did not know how to measure the dependent variable. This view is consistent with the analysis of their pretest IPRSs where it was noted that students found the identification of the variables to change and to measure difficult. For example some tried to measure the time for the reaction by plotting pH changes of milk rather than the time taken for a pH change.

Miss Mills: A common problem was the failure of students to identify **how** to measure the dependent variable. Some students didn't even realise that they had to measure it.

Mrs Grant added that the students did not conduct preliminary trials and that they did not understand how to use Universal indicator to measure the dependent variable. Her comment that students wrote down colour changes that didn't exist supports Assertion 7: Some students 'observe' changes that they think will occur in investigations rather than changes that actually occur. Also, she said that students did not take accurate measurements and tended to use non-standard measurements such as, "the test tube was one quarter-full", which provides further support for Black's (1990) assertion that some students don't perceive the need for quantification.

Mrs Grant: Some students didn't trial (conduct preliminary trials) at all. Some didn't understand the role of the indicator even though it had been explained to them. They seemed to be writing down changes in colours that didn't exist. I found it hard to work out what they were doing with the indicator. They didn't specify amounts accurately. They said that things were a quarter-full.

Mrs Cross said that students did not have a range of results and this comment was not surprising because in their previous school science, students had not been required to manipulate the independent variable or to gather sets of data. In addition, Miss Mills mentioned that students did not understand that the optimum conditions for the enzyme were at 37 °C in a slightly basic solution, and that they did not understand that they should only change one variable.

Miss Mills: Students didn't understand that they should be working with the optimum conditions and just vary one condition (variable).

In terms of *Processing data* teachers commented that students did not organise their data into tables or draw appropriate graphs. Mrs Cross commented on their poor graphing skills and that some students graphed pH instead of the time taken for the reaction. She added that they did not know what type of graph to draw, thus providing further support for research by Roth and McGinn (1997) and Assertion 17: Some students view continuous data as discrete data.

Mrs Cross: Their graphing was hopeless. Some tried to graph the pH change. They had no idea as to whether it was a line graph or a bar graph.

In summary, most of the teachers' comments about the pretest addressed students' abilities associated with *Conducting investigations* and to a lesser extent *Planning investigations* and *Processing data*. No teacher commented on the students' abilities at *Evaluating investigations*. This omission is consistent with questionnaire data gathered from the students and presented as Assertion 21: Students do not consider *Evaluating investigations* as learning about investigation competencies. Clearly both teachers and students considered the evaluation of the investigation of less importance than other investigation competencies. The following weaknesses in investigating competencies were stated by teachers during the May interviews.

Planning investigations: Most students' written plans lack detail and organisation, and failed to specify the amount of substances used.
Conducting investigations: Most students didn't conduct preliminary trials; they didn't understand how to measure the dependent variable; they used non-standard measures, they didn't control variables, they didn't understand the role of optimum conditions, and they wrote down colour changes that did not occur. This provides further support for Assertion 7 formulated in response to students' Investigation Planning and Report Sheets: Some students 'observe' changes that they think will occur rather than changes that actually occur.
Processing data: Most students didn't have enough data to graph, and they had problems drawing the graph and knowing what type of graph to draw. This provides support for Assertion 17: Some students view continuous data as discrete data.

Teachers' Perceptions of Students' Development of Investigation Competencies

At the outset, it needs to be stated that the teachers were pleased with their students' development of investigation competencies over the duration of the 30 lesson program. The comment by Mrs Cross after the final investigation, summarises teachers' perceptions of students' development.

Mrs Cross: They (students) have become so much better at doing the investigations but I don't think that they actually realise it.

The teachers said that some students did not perceive that they were learning a great deal from the investigations. They said that these students valued the acquisition of content knowledge and did not perceive the acquisition or development of skills and competencies, to be learning science. The comments by Mrs Cross and Mrs Grant support this view and are expressed in Assertion 36.

Mrs Cross: Some of them think they are wasting time. They view learning as the gaining of factual information. This could be our fault because we should test more science processes in our tests.

Mrs Grant: Some students are getting sick of it because they don't see it as learning. There is a lot of paper work and it is repetitive. Although some like Kara, really appreciate it. She can see the whole process and the rationale for it, so she is happy.

Assertion 36: Teachers consider that some students do not perceive developing investigation competencies to be learning science.

Although teachers' comments about students' development of investigation competencies were associated with specific investigations, it is possible to identify some common threads in their comments across the investigations, and these relate to the competencies of *Planning investigations, Conducting investigations, Processing data, Managing time* and *Being organised.* Most of comments that teachers made, however, were about the difficulties students experienced in *Conducting investigations*.

Planning investigations

The teachers said a major weakness in students' planning was their inability to plan appropriately and that this resulted in poorly presented written plans. This is consistent with the pretest data and may be due to students' difficulties in conceptualising the investigations. Students' poor planning ability is expressed as Assertion 37, and exemplified by Mrs Cross' comment.

Mrs Cross: A major student weakness was describing what was done. Diagrams were not used to help in explanations. The descriptions of collecting data were not done in a systematic way.

Assertion 37: Most students' written plans for investigations lack sufficient detail to describe the procedures they intend to follow.

Teachers (Mrs Cross and Mrs Grant) said that most students learned to use preliminary trials to familiarise themselves with the experimental procedures and to organise group work. Most students, added Mrs Cross, didn't use preliminary trials

to determine an appropriate range and/or interval for data collection. Teachers' views support Assertion 5: When planning investigations most students use preliminary trials to observe the reaction/experiment and to organise procedures. In addition they support Assertion 6: When planning investigations most students do not consider the range over which the data should be collected or the data collection interval.

Mrs Cross: Some of them don't associate trialing with trying to improve the experiment. Some don't distinguish between the trial and the actual measurements that they record. They just use it as a practice run.

Mrs Grant: Mostly they did preliminary trials to see how the equipment would work and to get themselves organised. They didn't have any idea about working out the range of readings or anything like that. (electromagnet investigation, November)

Conducting investigations

For some investigations, particularly the catalyst investigation and the electromagnet investigation, the teachers said that students found it difficult to work out what was the dependent variable and how it should be measured. Although the science curriculum had previously addressed the preparation of oxygen gas and the strength of electromagnets, students' abilities to access and link this knowledge to the investigations seemed to be critical factors influencing their understanding of how to proceed with the investigation. In contrast, the Panadol and ukulele investigations were simpler in science content and in terms of operationalising and measuring the dependent variable.

Teachers' comments associated with the measurement of the dependent variable are presented below. Comments about the conceptually simpler investigations from Mr Brogo (ukulele investigation), Mrs Cross (Panadol investigation) and Miss Mills (Panadol investigation), are contrasted against comments about the conceptually complex investigations from Mrs Cross (catalyst investigation) and Mrs Grant (electromagnet investigation). The comments

highlight the different difficulties associated with different investigations. Assertion 38 addresses these ideas.

Mr Brogo:	(Ukulele investigation) They handled it pretty well. It was easy and straight forward. They had learnt to do trials and ignore mistakes (musical notes that were not properly produced). They were quite well organised. No one seemed reluctant to use the computer and there was no difficulty in using the equipment.
Mrs Cross:	(Panadol investigation) They were much better organised. They found it more straight forward because they were just timing how long the Panadol took to dissolve. They come up with some weird ways of defining 'dissolving' because the last few specks take a long time to dissolve. Some were listening to the effervescence and were trying to time how long it took for the bubbles to stop popping.
Miss Mills:	(Panadol investigation) Some seemed to like this investigation because they knew that their data were reasonably accurate. Others thought that it was beneath them and too trivial. It was certainly easier that the other investigations. It probably should be earlier in the sequence of the investigations.
Mrs Cross:	(Catalyst investigation, May) They had trouble trying to work out how to measure the dependent variable. (Catalyst investigation, August) It went well except that they had no idea how to collect the oxygen. I had to refer them back to when we had collected oxygen before. They don't seem to associate their class work with the investigations. They still had trouble measuring the dependent variable. Some chose inappropriate measures such as when the first bubble appeared. Most worked out that they needed a line on the test tube. Virtually all did (the effect of changing) the amount of catalyst.
Mrs Grant:	(Electromagnet investigation) It was more difficult than the ukulele investigation. Students had to sort out the experimental set up, the manipulation of the equipment **and** the measurement side.

Assertion 38: Students' familiarity with the dependent variable, and how it is operationalised and measured affects the difficulty of investigations.

Interestingly, the ease of manipulating the independent variable was not identified by teachers as a factor contributing to the difficulty of the investigation, even though studies in the United Kingdom have shown that students find continuous independent variables difficult to operationalise (Duggan et al. 1996). Teachers, however, did identify the setting up of equipment as a factor influencing

difficulty, Assertion 39. This is implied in previous comments by Mrs Cross about the catalyst investigation. She said students had trouble working out how to measure the dependent variable and how to collect oxygen gas, even though they had studied the collection of the gas a few weeks earlier. She added that the equipment needed to be set-up to collect the gas and that this created a stumbling block for some students. Mrs Grant's following comment about the electromagnet investigation also draws attention to the complexity of setting up equipment. Perhaps this additional element in conducting investigations is sufficient to create an information overload situation for some students (Johnstone & Wham, 1982).

Mrs Grant: The students came up with the variables to investigate. Some of them had bits of the investigation correct but they found it harder than the ukulele investigation because they had to set up the circuit. There were many more decisions to make. I don't think that it was too difficult. (electromagnet investigation, November)

Assertion 39: Students find investigations that involve setting up equipment more difficult that those that do not require equipment to be set-up.

Assertions 38 and 39 also imply that investigations vary in their degree and nature of difficulty. These data support notions of task and domain specificity associated with performance tasks (Gipps, 1994) and the impact of contextual factors on performance.

Teachers commented that from the first few investigations students learned to repeat trials and to average results. The following comments by Mr Brogo and Miss Mills indicate this, and are further support for Assertion 8: Most students learn to repeat or replicate the data collection and to average the results.

Mr Brogo: The trialing was a bit rushed by the students. They had learnt to do trials and ignore mistakes (frequencies produced by irregular plucking of the ukulele string). They were quite well organised. They seemed to have grasped the basics of repeated trials. They

seemed to understand to take about five sets of data whereas in the past they would only take three sets of data. They also repeated trials and averaged results. (ukulele investigation, August)

Miss Mills: The students seemed to have got the idea about repeating trials and they managed to collect quite a lot of data. (Panadol investigation, November)

Processing data

During May and August, teachers commented on students' limited competencies at processing data, particularly their poor graphing skills. However, by the end of the program they said that they were pleased with the students' improvements in this area.

Mrs Cross They were poor at drawing tables and not all students got around to doing the graph. They need a bit more on graphing skills. A lot of them don't know to put the independent variable on the horizontal axis. (catalyst investigation, May)

Managing time and being organised

At the beginning of the program teachers commented on the students' poor organisational skills, however, they said that after a few investigations the students realised the need to become more organised in order to complete the investigations on time. Mrs Cross' comment below exemplifies how some groups functioned and Mrs Grant said that some groups did not work well. All of the teachers thought that students had sufficient time to complete the investigations but the students complained that they did not. This is exemplified by Miss Mills' comment.

Mrs Cross: Often one student in the group tried to do everything at once. Later it might be another student. For example they would try to put in the stopper, start the timer and put the gas delivery tube in the right place. They weren't good at sharing the organisation of tasks. (catalyst investigation, May)

Mrs Grant: Some students (Kate and Daisy) had problems working in the group and didn't work well. (electromagnet investigation, August)

Miss Mills: The students felt that they didn't have enough time but I thought that they did. They seemed to enjoy it. They said they were rushed but they seem to say this whatever they are doing. (catalyst investigation, May)

Aspects of the Cognitive Apprenticeship Model of Instruction

Teacher Modelled Investigations

The teachers were briefed about the concept of modelling and its role in the cognitive apprenticeship model of instruction. They modelled investigations to the whole class, similar to the concept of "global modeling" described by Javela (1996, p. 110). Each teacher had previously trialed aspects of the data gathering procedures before the lesson.

Teacher modelled acid and carbonate investigation

The acid and carbonate investigation was performed by Mrs Cross (Class TN) and Miss Mills (Class TC) during the first sequence of lessons (May) and by Mrs Cross (Class SC) during the second sequence of lessons (August).

The general consensus of the teachers was that it was too difficult for students to observe the investigation. Also, they said that they found it too "fiddly" for a demonstration and "too time consuming gathering the data." They added that they tried to present too much information and, in order to get through the work, they talked too much. As a consequence, students lost interest because they were not sufficiently involved. The comments below support these interpretations.

Miss Mills: The lesson did not involve students in sufficient activity. There was not enough for them to do. I found it far too teacher centred and quite frustrating. I attempted to convey too much information and it was very time consuming. I think the students lost interest.

Mrs Cross: I felt too rushed.

During August Mrs Cross expressed similar views to those she expressed during May. However, when asked what the students learned from the investigation she said that it was "good" for the revision of terminology.

Mrs Cross: I don't really like it because it's too fiddly and the students get bored and chatty while it is being set up. At low concentrations the reaction is too slow. The Panadol investigation would be easier to demonstrate. It's good for the revision of terminology. They seemed to have forgotten everything anyway. Probably its good to show the importance of trialing and taking repeated measurements. With the setting up of the equipment the students ask continually how to draw it. It would work better if it was easier to set up. I wanted to get heaps of trials done and the graph drawn but I didn't get to the graph. If it was simpler they would get more out of it. Not all students were attentive. It would be better to get them to do it and to go through it with them in a stepwise fashion.

Teacher modelled pitch of a closed pipe investigation

The pitch of a closed pipe investigation was modelled by Mrs Grant (Class SC) during May, and Mrs Grant (Class TN) and Mr Brogo (Class TC) during August. The set-up of the equipment for this investigation was far more simple than for the acid and carbonate investigation. It was only necessary to change the independent variable, the depth of water in a bottle, because the computer measured the dependent variable and students read the data from the screen. This may be a reason why these teachers were more positive about the value of modelled investigations. One of the main teaching points was the use of the computer software because students needed this information to perform the ukulele investigation. Therefore, the teachers modelled the investigation process as well as specific data collection techniques that were needed for the next investigation.

Mrs Grant acknowledged that the modelled pitch of a closed pipe investigation had value because it was closely related to the ukulele investigation, however, she stated that her first class (Class SC, May) became a little bored because it took too long. She added that she thought it would be easier to demonstrate to smaller groups because not all student could see the computer.

Mrs Grant: I can see the value in the activity because it helps with the ukulele investigation. The students got a bit bored. It took a long time and they stopped listening. I think that, perhaps it would be better to model to smaller groups because it's hard for them to see the computer.

For the second sequence of lessons during August, when Mrs Grant was asked, "How did it go?" she replied "Good! They stuck with it well." She added that it was difficult for students to see because there were too many around the computer. Mr Brogo said that it went "OK!" He stated that the modelled investigation prepared them well for the ukulele investigation that was to follow.

Mrs Grant: There were too many around the computer. Next year I will divide the class in two so that one half does the worksheet while the others work on the computer.

Mr Brogo: It focused them quite well. Almost all groups chose length and frequency for the ukulele investigation, therefore it was compatible with the ukulele investigation. It took a long time to get through the work.

From the discussion about the teacher modelled acid and carbonate investigation and the teacher modelled pitch of a closed pipe investigation it is evident that the acid and carbonate investigation was more difficult to model. The teachers commented about the length of time taken for the investigation, problems that students had in observing the investigation, and that the acid and carbonate investigation was difficult to set up. Assertion 40 has been formulated from these data.

Assertion 40: Teachers perceive that teacher modelled investigations are difficult to implement in whole-class settings because

(a) they are time consuming and students become off-task,

(b) sometimes not all students can observe the data collection, and

(c) setting up the equipment can be fiddly.

Modelling, from the teacher's perspective, and **observing,** from the student's perspective, are identified by Collins, Brown and Newman, (1989) as central to the cognitive apprenticeship model of instruction. However, because of the difficulties

associated with modelling investigations in whole-class settings, the teachers believed that it would be better for students to perform the investigation and be provided with help from the teacher when needed. These ideas are consistent with the cognitive apprenticeship notions of providing coaching, guidance and scaffolding for the learner as they execute or practise a new skill. Also, they are likely to be consistent with the traditional apprenticeship model of instruction in which an apprentice would start work and request help from a master craftsman when needed. Comments by Miss Mills and Mrs Cross support the view that it would be better to provide help when necessary rather that try to model an entire investigation.

Miss Mills: I suspect that students would learn more through doing an investigation with the teacher's help.

Mrs Cross: It would be better to get them to do it (an investigation) and for us to go over it in a stepwise fashion.

These comments support students' views collated from the questionnaire data (Assertions 28 and 29). Clearly both staff and students have indicated that modelling an entire investigation is not an ideal teaching strategy and Assertion 41 is based on teacher recommendations and student questionnaire data. The ideas expressed in this assertion are consistent with "situation specific modeling" described by Javela (1996, p. 100). She contends that situation specific modelling is more directed at students' problems; more likely to develop further student inquiry, and more reciprocal which implies that the teacher-students interactions are more personal.

Assertion 41: Teacher modelling of investigation competencies may best be implemented with small groups of students so that teacher guidance can be

(a) provided when needed by the group,

(b) highly focused on the difficulties encountered by the group, and

(c) at a personal level.

Teachers also expressed the view that they tried to achieve too much during the modelled investigation. The range of investigation competencies that could be developed include *Planning investigations*, *Conducting investigations*, *Processing data*, and *Evaluating investigations*, as well as *Writing a report*, *Working cooperatively*, *Attending to detail*, *Managing time* and *Being organised*. Therefore, it is suggested that if teachers wish to develop students' investigation competencies then they should select a small number of competencies and model these competencies. Students' motivation for learning investigation competencies is likely to be higher if they are aware that the competencies are required to perform the next investigation. For example, the data gathering skills modelled in the pitch of a close pipe investigation were needed for the ukulele investigation. Hence, this may have contributed to students rating this teacher modelled investigation above the acid and carbonate investigation in terms of contributing to learning investigation competencies (Table 7, p. 110). The notion of modelling a small number of investigation competencies at one time is supported by comments from Mrs Grant and summarised in Assertion 42.

Mrs Grant: I think it would be better to identify the skills that students need to develop, and to model smaller chunks like data collection, graphing, ways to measure the dependent variable and so on.

Assertion 42: The focus of teacher modelled investigations should be on developing a small number of competencies.

Teacher Guidance: Working in the Zone of Proximal Development

Teachers acknowledged that the investigations posed students many problems. Mrs Grant and Mr Brogo said that many students didn't know how to proceed when faced with a problem and that sometimes this resulted in off-task behaviour. As a consequence, classroom management difficulties may arise, Assertion 43. When faced with difficulties students adopted different behaviours; some asked the teacher for help, some went around the room trying to see what the other students were doing, and some socialised. It may have been that the difficulty of the investigation was inappropriate for some students because it was not in their zone of proximal development (Vygotsky, 1986). One way to address this problem was suggested by Mines (1995). He advocated that teachers use checklists to determine whether students have the necessary knowledge and skills, and know what to change and measure **prior** to an investigation if the investigation is to be used for assessment.

Mrs Grant: Classroom management was quite difficult. There was a lot of movement in the room. It was difficult to keep them on task. Some students 'sit' on the problem. They don't naturally engage in problem solving and they stop at this point. There is a problem with what to do with students when they encounter this problem. Some groups really get into it. They then spend more time on it than the others. Some engaged at a lower level and had 'free' time to wander around. (electromagnet investigation, August)

Mr Brogo: I heard some kids say, "This hurts the brain, it's too difficult." They definitely have to think things out. Some kids opted out of the problem. They started to mess around. Some wanted me to set up the circuit for them. (electromagnet investigation, August)

Assertion 43: Classroom management problems can arise when students find investigations too challenging.

Teachers viewed their roles differently in helping students to solve problems associated with the investigations. Mr Brogo tried to get student to work through their own problems as indicated by his comment that, "Some wanted me to set up

the circuit for them," and also by his comment below. In contrast, Mrs Grant was more comfortable helping students so that things turned out right. These differing views expressed by teachers are presented as Assertion 44. This assertion is complementary to students' perceptions of different teaching styles. Some students prefer teachers who give full explanations and some prefer teachers who encourage and assist them to solve their own problems, Assertion 27. Maintaining a balance between providing suitable guidance and leaving sufficient scope for students to think independently has previously been identified by Mines (1995, p.14) as dependent on skilful questioning.

Mr Brogo: They don't naturally want to work through things but I think its better if they work things out for themselves. (electromagnet investigation, August)

Mrs Grant: Students adopted many different approaches and I was not sure of how much guidance that I should give. I didn't help them and wished that I had. This caused some conflict. Some ran off the rails with their design and I found that distressing.

Assertion 44: Teachers have different views on how much guidance to give students performing investigations.

Scaffolding and Fading

A scaffolded Investigation Planning and Report Sheet (Appendix F) was provided to students for each investigation except the pretest and posttest investigations (Appendix D) where less scaffolding was used. The posttest represented 'fading' and the pretest matched this format so that comparisons could be made between the students' pretest and posttest performances. Teachers' commented about the structure and questions on the IPRS rather than their purpose and the way they were used. This is illustrated by the following comments. No assertions were formulated from these data.

Mrs Grant: The worksheet could be structured a little differently. It seems as if they ask the same thing on different pages.

Mrs Cross: Some students planned it methodically and then had to change their plan after they had played with the equipment. Perhaps before planning they need to play and choose the variables. They would then think more about conducting the investigation. Then they could write the plan. There is not enough space for the preliminary trials if they decide to change their plan. They then have to write a new hypothesis. Maybe the hypothesis writing should be later. About five students were messed-up with the write-up because they had to change their plan. Maybe make the plan a 'rough plan'.

Self-Reflective and Metacognitive Skills

The students participated in three types of activities that involved self-reflective and metacognitive practices. These were the worksheet on marking an investigation, the questionnaires, and the evaluation section of the Investigation Planning and Report Sheets. The teachers made inferences about students' self-reflective and metacognitive abilities.

The worksheet on marking an investigation (Appendix G) was designed to encourage students to evaluate an investigation that was performed and written-up by the preceding Year 9 students. The investigation had specific weaknesses. These included graphing, averaging data, selecting an appropriate sample size, drawing conclusions that were consistent with the data, and controlling variables.

Mr Brogo: The students found it very interesting. A few didn't understand why they were doing it. It involved a lot of metacognition and I think their abilities to do that are based on their level of maturation. I think it was good for making comparisons with their own work.

Mrs Cross: It was the first time that they had done anything like this and at first they didn't see the value in it. I think that by the end of the lesson most had identified mistakes that the other students had made, particularly with the multitude of graphs. They had to think quite a bit. They are used to accepting information and not thinking about whether things are right or wrong.

To complete the questionnaires and the evaluation section of the IPRS students had to reflect on, and document what they had learned about doing the investigations. The teachers' comments indicated that students didn't see the value in thinking about their learning and Miss Mills stated that at Year 9 she believed that most of the students were insufficiently mature to think about how and what they learned. Teachers' perceptions of students not being sufficiently mature to reflect on their learning is expressed as Assertion 45.

Mrs Grant: Generally they don't like doing it. They say, "Why do we have to do this? They tend to write the same things down for all answers. Most of them don't know what they have learnt. They just aren't able to pinpoint areas.

Miss Mills: They seemed to forget what they learn very quickly. They (questionnaires) helped to remind them about their learning. Some of them looked at their worksheets (IPRS) to answer the questions. I think that by the last one they were a bit sick of it. They don't see that this sort of thing helps them. They just see it as ng something for someone else. I think that they're a bit too imma..re to think about how and what they learnt.

Assertion 45: Teachers believe that Year 9 students are not sufficiently mature to reflect on their learning.

This teacher belief is not supported by research findings. For example, Baird (1986a, 1986b) conducted a six month action-research study and found that when Year 9 and Year 11 students applied evaluative cognitive strategies during lessons they became more informed, purposeful learners and exercised greater control over their learning. It may be that in this study students were not provided with sufficient opportunities to practise and develop appropriate metacognitive skills and because of this teachers believed that the students were too immature.

Assessment

The three classes in the study experienced different assessment regimes. Class TN had their investigations assessed by a teacher and their performances were ranked and graded relative to other students in the class. Class TC had their investigations assessed by a teacher and they received a grade for criteria relating to the investigation. Students in Class SC worked in groups to assess their own investigations from a master answer sheet and graded their work according to the same criteria as Class TC. Details of the assessment regimes are in Appendix I.

Teacher Assessed Norm Referenced Assessment: Class TN

The teachers of Class TN said that it was difficult to determine an order of achievement and that the assessment took a long time. These points of view have been expressed respectively as Assertions 46 and 47. One possible reason for the difficulty in ranking students' performances may have been because the classes were selected from a narrow ability range of students in the Year 9 cohort. A high ability and a low ability class did not participate in the program and the three participating classes comprised middle ability students. Another reason may have been because students worked together for the planning and conducting phases of the investigation and only completed the write-up of the investigation by themselves. As a consequence, they would have shared ideas before the write-up and this may have resulted in less variation in their Investigation Planning and Report Sheets.

Mrs Cross: I marked them in groups because it took too long to do it individually and the groups tended to write the same things. I put them in an order and then assigned the grades. I was not sure about the order (rank) because they seemed to be about the same standard. Because I had marked them quickly and superficially I didn't give as much help to individual groups as I would have liked. I went through each section with the whole class on the whiteboard. I felt that they were interested in writing down the corrections. (May)

Mrs Grant: Marking was a problem. I couldn't scan their work and make a judgement. Their work was too similar. It was impossible to

separate them to norm reference them. I didn't like the norm reference marking. It seemed as if they weren't as bad as a C+ or a C. I said (to them) that it was very hard to get an A. (August)

Mrs Cross: It was OK. It takes a long time to sort them and sometimes it's hard to order them, especially for kids in the middle. Afterwards I went over quite a bit with them on the board (whiteboard). (November)

Comments about the difficulty in ranking students' achievements to fit a normal distribution support a documented weakness of norm referenced assessment (Gipps, 1994). Clearly for this assessment procedure to compare students, test items need to have a high degree of discrimination. It is likely that the practice investigations that were not conducted under test conditions, were not sufficiently discriminating to achieve this process easily and quickly.

Assertion 46: Teachers found it difficult to rank students' Investigation Planning and Report Sheets in order of achievement.

Assertion 47: Teachers found norm referenced assessment of the investigations to be very time consuming.

Teacher Assessed Criterion Referenced Assessment: Class TC

The teachers of Class TC found the criterion referenced assessment of students' IPRSs very time consuming. To speed up the process they sorted the students' IPRS into the groups in which they worked. They then assigned grades to the individuals in each group, addressing the criteria one at a time. These findings are presented as Assertions 48 and 49 and are supported by the comments of Miss Mills and Mr Brogo. Sorting students' IPRS into the groups in which they worked, prior to the assessment, was also done by Mrs Cross when she assessed the IPRSs of Class TN.

Miss Mills: The marking took a long time. Their reports were long and you had to flick through a lot of pages. I was worried about the consistency of my marking. The groups tended to write the same thing. In the end I sorted the students into groups and virtually group marked. It took a long time writing comments. In the end I only wrote brief comments and then went over things with the whole class.

Mr Brogo: It was easier to mark in groups. I did not find this difficult. I like working with descriptors and I am familiar with this. I made comments on the worksheets. I think there was an improvement in the method they used in the second investigation (electromagnet investigation), particularly in their description of the results. I' m not sure what it was due to. Students were keen to see their marks. They didn't initiate any discussion about their grades. I would have expected some discrepancy between the teacher's marking and the students' expectations. I did go through the investigation afterwards. Some students improved on the descriptive component.

Miss Mills: The kids liked tallying up their As, Bs and Cs. They quite like this marking. It gives a good indication of where to improve but it doesn't really tell them how to improve. It takes too long to write out explanations for all their mistakes. I think that in terms of time it's better to do it in class (in a whole-class setting). The trouble is that mostly they're not interested and don't listen when you go over it because it's finished.

Assertion 48: Teachers found criterion referenced assessment of the investigations to be very time consuming.

Assertion 49: Teachers found that with criterion referenced assessment it was easier to assign grades for individual students group by group.

Student Assessed Criterion Referenced Assessment: Class SC

The teachers stated that they did not like this form of assessment because they did not get sufficient feedback on students' performances. Hence, they believed that they were unable to give appropriate guidance for future learning. This is supported by Assertion 50 and the comments by Mrs Grant and Mrs Cross.

Mrs Grant: I didn't really like it when the students marked their own work. I didn't know how they went. There was no feedback to me. I couldn't help them afterwards because I didn't get the feeling for what they did

wrong. I did help some individually though, when they asked questions. I don't think that they got much out of it. Most of them are not very focused on improving their work. They just want to get things done, especially if marks don't count for a grade or anything. (ukulele investigation, May)

Mrs Cross: Because students went over the assessment themselves I didn't feel that I went over it enough. I don't think that some of the points were addressed sufficiently in the worked solution. (catalyst and Panadol investigations, August)

Mrs Grant: I don't like marking this way. I don't know how much the kids get out of it. Not very much I think. The trouble is that I don't pick up on what they can't do so it's hard to help them. (electromagnet investigation, November)

Assertion 50: Teachers do not like student self-assessment because they do not get feedback on students' performances and as a consequence they find it difficult to address students' errors.

This assertion has implications for the effectiveness of formative assessment during laboratory investigations. Clearly, as students performed the investigations teachers did not gather sufficient information through observation or questioning to be able to make informed judgements to improve the teaching and learning program.

Summary of the Chapter

This Chapter has addressed data gathered from teacher interviews conducted on four occasions. From the data 15 assertions were formulated and these are presented in Figure 19. Four assertions relate to students' investigation competencies. Assertion 37 referred to students' lack of attention to detail in their planning and may be due to their failure to conceptualise the whole investigation. Assertions 38 and 39 relate to factors affecting the difficulty of investigations and these support the notions of task and domain specificity associated with performance tasks.

Six assertions related to the cognitive apprenticeship model of instruction. Teachers' perceptions of the modelled investigations in a whole-class setting were not positive (Assertion 40). The need to pitch investigations in the zone of proximal development was borne out by classroom management problems arising when the tasks were too difficult (Assertion 43).

Five assertions were formulated from teachers' perceptions of the assessment regimes. Some were associated with the time taken to assess the investigations (Assertions 47 & 48) and the problem of not getting feedback on students' performance when students assessed their own investigations (Assertion 50).

Support for previously documented data and assertions emerged. For example, teachers confirmed pretest data from the IPRSs indicating that students tried to measure a dependent variable that was different from that which was intended, and also that students' graphing skills on the pretest were poor. In addition, support for Assertions 5, 6, 7, 8 and 17 from Chapter 4 and Assertions 21, 27, 28 and 29 from Chapter 5 was identified.

Assertions

Investigation competencies

36 Teachers consider that some students do not perceive developing investigation competencies to be learning science

37 Most students' written plans for investigations lack sufficient detail to describe the procedures they intend to follow

38 Students' familiarity with the independent variable, and how it is operationalised and measured affects the difficulty of investigations

39 Students find investigations that involve setting up equipment more difficult than those that do not require equipment to be set-up

The cognitive apprenticeship model of instruction

(a) Teacher modelled investigations

40 Teachers perceive that teacher modelled investigations are difficult to implement in whole-class settings because
(a) they are time consuming and students become off task,
(b) sometimes not all students can observe the data collection, and
(c) setting up the equipment can be fiddly

41 Teacher modelling of investigation competencies may best be implemented with small groups of students so that teacher guidance can be
(a) provided when needed by the group,
(b) highly focused on the difficulties encountered by group, and
(c) at a personal level.

42 The focus of teacher modelled investigations should be on developing a small number of competencies.

(b) Teacher guidance

43 Classroom management problems can arise when students find investigations too challenging.

44 Teachers have different views on how much guidance to give students performing investigations.

(c) Self-reflective and metacognitive skills

45 Teachers believe that Year 9 students are not sufficiently mature to reflect on their learning.

Assessment

(a) Teacher assessed norm referenced

46 Teachers found it difficult to rank students' Investigation Planning and Report Sheets in order of achievement.

47 Teachers found norm referenced assessment of the investigations to be very time consuming.

(b) Teacher assessed criterion referenced

48 Teachers found criterion referenced assessment of the investigations to be very time consuming.

49 Teachers found that with criterion referenced assessment it was easier to assign grades of individual students group by group.

(c) Student assessed criterion referenced

50 Teachers do not like student self-assessment because they do not get feedback on students' performances and as a consequence they find it difficult to address students' errors.

Figure 19. Summary of assertions developed from teacher interviews

CHAPTER 7

STUDENT INTERVIEWS

Overview of the Chapter

This Chapter addresses the perceptions of three groups of students about the main themes of the research; the acquisition of investigation competencies, aspects of the cognitive apprenticeship model of instruction, and the different assessment regimes implemented in the study. Groups of three students were interviewed within two days of finishing the 10 lesson sequences during May, August and November. Group TN (Kim, Tammy and Gay) was from Class TN; Group TC (Jessica, Olive and Gemma) was from Class TC and Group SC (Christie, Jody and Harue) was from Class SC. The same groups of students were audio and video recorded as they performed the investigations and data associated with these recordings are presented in the following chapter. The interviews were semi-structured and broadly addressed the themes of the research. They were of approximately 20 minutes duration and recorded on audio and video tapes and later transcribed. The data augment student questionnaire data (Chapter 5) and provide a complementary perspective to the teacher interview data (Chapter 6). Some additional assertions have been made, and also the data confirm a number of assertions previously generated. In addition, the data present a contrasting perspective to one of the assertions previously generated.

The investigation competencies section deals with group impressions of the pretest and their perceptions of the competencies they developed. The discussion about aspects of the cognitive apprenticeship model of instruction addresses teacher modelled investigations, coaching and articulating, scaffolding and fading, and self-

reflective and metacognitive skills. Finally, the Chapter discusses students' perceptions of the assessment regimes used in their classes.

Investigation Competencies

Students' Perceptions of Investigation Competencies from the Pretest

Data from the student interviews did not result in the formulation of any new assertions regarding performance on the pretests, however, there was considerable support for previous assertions. The three groups did not perform the intended investigation, factors affecting the time taken for lipase to react with milk, but investigated the factors affecting the amount of fatty acids produced by milk. Students' tendencies to investigate this alternative problem were documented previously in Chapter 4, from the analyses of students' Investigation Planning and Report Sheets, and Chapter 6, the teacher interviews.

Group TN said they investigated the "Temperature and the pH numbers," and on their IPRS they wrote an hypothesis that was unlikely to be supported by data (Assertion 1), "The higher the temp. of the water the more acidic milk becomes in the presence of lipase." This indicates that they thought an increase in temperature would increase the reactivity of the enzyme and hence the amount of acid produced. This assumption is consistent with Assertion 2, in which students related an increase in temperature to a decrease in the time for the reaction. A description of Group TN's IPRS was presented previously (p. 96) and it describes how Kim, Gay and Tammy rejected their hypothesis and proposed a new conclusion based on insufficient evidence (Assertion 14). During the interview they commented, 'Ours didn't work out like everyone else. Ours didn't change (colour). We added more indicator and timed for the same time (for each test)." It was evident that there was no pH change and that their results were their initial pH readings. Based on these

data the students concluded that their "hypothesis was not right" because the temperature effect was not as they had hypothesised. It is sobering to note that on Tammy's questionnaire, she wrote, "Temperature doesn't affect the breaking down of fatty acids in milk in the presence of lipase." From their data the group reached a conclusion that is not consistent with established scientific theory. It is somewhat heartening that they did not shape their data to support their original hypothesis, but disheartening that Tammy's conceptual understandings are not consistent with established theory. This example shows that flawed data gathering procedures may result in the development of conceptual understandings that are inconsistent with those of scientists.

Group TC proposed to investigate "the more milk the higher the pH." They varied the amount of milk (2 mL, 6 mL and 10 mL) but did not control for the total volume of the solution (Assertion 4). In effect they diluted the enzyme (lipase) by using more milk so their hypothesis was likely to be supported. "I didn't really understand it," says Jessica. "Wasn't it the higher the pH the fattier the milk?" The group discussed this and agreed that they changed the amount of milk and tried to measure the pH change. "It was hard trying to get the pH colours (from the chart) to match up (with the milk). We got confused. The conclusion wasn't what we thought. We couldn't get a good colour change. The changes were about a half of a pH." Changes of half a pH unit are not indicated on the pH chart and this response supports Assertion 7: Some students 'observe' changes they think will happen in reactions rather than changes that actually happen. Their interview data was then matched with their IPRS on which they each had written, "... and we discussed the colour change and we got an all round decision on the colour." Jessica and Gemma had no data (colour changes) recorded on their IPRS and as a result did not complete the last two sections. Olive completed the IPRS and without indicating whether the group had controlled for time she wrote, "Sample A went from pH 6 to pH 4. B went from pH 6 to pH 5. C went from pH 6 to pH 6 (sic)." Hence she achieved an

increase in the final pHs from four to six with dubious colour changes that corresponded to an increased volume of milk. Olive also included a bar graph for the pH changes of the three milk samples (A, B and C), however, the vertical pH scale on her graph was not meaningful. She concluded the inverse of the groups' hypothesis by recording, "The less milk the lower the pH value." From the IPRS it seemed that Olive had 'confirmed' the hypothesis, with fudged data thereby adding support to Assertion 12: Some students confirm their hypothesis with data from dubious sources. These findings are also consistent with previous research on students fudging data (Rigano & Ritchie, 1995).

Group SC hypothesised, "The more fat the higher the acidic rating," and afterwards commented, "We should have done what the worksheet (IPRS) said. I don't think we planned it very well. We should have done more preliminary trials. We should've done better, and planned better." They also recognised that they had not collected their data under optimal temperature conditions and this supports Miss Mills' comment (p. 154). We should have "done it (the investigation) in a hot bath." In addition, Group SC said that their measurements were not very accurate because, "It was hard to get colours (colour changes). There are many colours in between and things on the chart (pH chart for Universal indicator colours). Some colours didn't relate to what we got in the milk. It was not very accurate because we had to add heaps of indicator to see the colours." Even though these students highlighted problems in identifying the pH, according to their IPRS they managed to collect data that supported their hypothesis and were "very" confident in their results. When asked how they could improve on the investigation they replied that it was important to "hypothesise properly." This is likely to mean it is important to select an hypothesis that is supported by data. They added that they needed "to follow the instructions on the sheet," and that "they should have done more trials before getting started." They said, "Our results matched the hypothesis but it was not really what we were meant to do. We did what we set out to do but we got it wrong." These

responses indicate that students are attempting to achieve some preconceived right answer, behaviours fostered by previous experiences with traditional laboratory exercises (Fairbrother & Hackling, 1997).

Each group had difficulties determining how and when to measure the dependent variable. With their hypotheses they could have controlled for time and judged the colours after a fixed time period or they could have controlled for indicator colour and timed how long it took for the indicator to change to a predetermined colour. Difficulties associated with measuring the dependent variable have been discussed previously in relation to the IPRS in Chapter 4, and the teacher interviews in Chapter 6. These findings parallel those by Duggan et al. (1996) indicating that students experience difficulties when defining and operationalising a continuous independent variable. Although no new assertions have been formulated from student interview data about pretest performance, confirming evidence for discussions in previous chapters has been presented. In particular, students' testimonies have supported Assertions 1, 2, 4, 7, 12 and 14.

Students' Perceptions of Developing Investigation Competencies

The groups were asked what they had learned from each investigation they performed following the lessons during May, August and November. Their comments and the investigation competencies they identified, were similar to the written responses to the open-ended questions on the questionnaire (Assertion 22). They said they had improved at *Planning investigations*; *Conducting investigations* including preliminary trialing, replicating or repeating trialing and controlling variables; *Processing data* including recording results in tables, averaging and graphing; and had learned that when *Evaluating investigations* an hypothesis is supported but not proven. In addition, the groups said that they had improved at completing the IPRS, and knowing what the teacher wanted written on the IPRS.

Time management continued to be a problem as indicated in a comment from Group SC. The following comments were made by the groups after the posttest investigation and are associated with investigation competencies.

Group TN: It's been OK! Good! I think we've done too many. I think about three is enough. We learnt about the procedures to follow. We got faster at **planning**, more efficient. We got better at **controlling variables, conducting trials** and making sure that we don't run out of **time**. We got better at recording our results in a **table** and doing a **graph**. We learnt how to do graphs with their axes and stuff. We're better at it now. I still don't really get when to do bar graphs and line graphs (Kim). We learnt that the **hypothesis should be supported** rather than proven. We learnt that you have to **answer the last page**. It's better than normal work (the usual lessons). The other prac. work is good, but it's a lot shorter. We liked cutting up the kidney. We like doing experiments but we don't like writing them up. It's (writing up) a bit boring. We like the practical side.

Group TC: Compared with the other work it's more interesting because we did it. We all get different results. We don't have to learn stuff. We learnt how to do the investigations. The **preliminary trials** are a bit repetitive. With the preliminary trials you find out what things might affect your results. We learnt that you have to do **three experiments** and find an **average**. You can **never be certain of your conclusions**. You have to put units at the top of the **table** and use headings. We know when to do a bar **graph** or a line graph. We have improved at setting out and **knowing what the teacher wants**.

Group SC: Apart from the questionnaire the worst thing is the amount of **time.** It's hard to get into it at the beginning and then at the end you're rushing to get through it. People would really benefit from a longer lesson. It would be good if we had a double period.

Group SC's comment about requiring more time is in conflict with teachers' views which were expressed in Chapter 6. Teachers believed that students had sufficient time but that their organisational skills could be improved. It is also evident from the preceding comments that there is considerable support for Assertion 22: Investigation competencies that students claim to have learned include (a) planning investigations that comprises writing hypotheses, identifying variables, overall planning and preliminary trialing that can result in modifications to the plan (b) conducting investigations that comprises using equipment, controlling variables, repeating trials, being accurate and working safely (c) processing data that comprises

using tables, averaging results, using graphs and drawing conclusion (d) writing reports (e) working cooperatively (f) attending to detail (g) managing time and (h) being organised. The questionnaire data indicated that students did not identify *Evaluating investigations* as learning about investigation competencies (Assertion 21). Although the students who were interviewed mentioned that they had learned some of the finer points associated with *Processing data,* such as hypotheses can be supported but not proven and that conclusions need to be made with care, they did not suggest specific ways to improve their results or to evaluate the investigation. Therefore the interview data also support Assertion 21.

Assertion 38: Students' familiarity with the dependent variable, and how it is operationalised and measured affects the difficulty of investigations, was supported for example when Group TN commented, "It was easier to measure (the dependent variable) because the computer did it." Assertion 8 was supported when students referred to the fact that they had learned to carry out repeat trials. For example, Group SC said, "We learnt that you had to do lots of trials and to be as accurate as possible with everything." Similarly Group TC said, "We did it about three times." Group TN said that they found planning difficult and this matches the teachers' perceptions that students had poor planning skills, Assertion 37, and is consistent with previous research findings in Western Australia (Hackling & Garnett, 1995).

During the interviews it emerged that some students were aware of differences between their **normal** science lessons and science investigations, and this is illustrated by Harue's comment. The fact that Harue believed that they were "not doing anything" supports Assertion 36, that the teachers believed that some students did not perceive developing investigation competencies to be learning.

Harue: Sometimes I feel as though we're not doing anything. For the other work we get notes to tell you what to do. For the investigations we had to learn as we went along. We didn't have many instructions. I prefer the other booklets (normal science) because I've got something

> to study and if I don't understand it I can go back and redo it. Whereas in this one (investigations) if I don't understand, it doesn't make any difference because I can't go back and find the answer. I have to do it **all** again.

Comments the groups made about their organisational skills support comments made by Mrs Cross (p. 160). She said that students did not get themselves organised and that often one person tried to do everything. In comparison, Group TC said, "To start with we didn't have our own jobs and we kept running everywhere. Next time we would have our own jobs and we need more time (another period). Group TN commented, "Getting organised we had to cooperate in groups, not scream at each other. Some groups got mad with each other and didn't want to do things. We got quicker at getting the equipment."

The groups preferred science investigations in which they investigated problems for which they did not know the answer and this notion is expressed as Assertion 51. It seemed as though they wanted to solve a science problem that was a real problem for them. There was almost sense of disappointment and futility associated with the investigation when they believed that they already knew the outcome. They perceived that the results of the Panadol investigation and the ukulele investigation were "general knowledge." The comment by Group TN revealed that they did not like the electromagnet investigation because Mrs Grant had previously shown them an electromagnet and they knew what the results would be. Group SC said that they valued the lipase investigation more highly because they learnt about the reaction of lipase and milk. It was also the first investigation that the students performed.

Group TN: We didn't like it because the teacher had already done it so we knew what was going to happen. (The comment was made by Olive although Gemma said that she couldn't remember doing it before.) We worked better. We don't know why. We knew what the results would be. (electromagnet investigation, August)

Group SC: The lipase was the most valuable because we learnt everything about it (the reaction of lipase with milk). Whereas the other one we knew what would happen. (ukulele investigation, May)

Assertion 51: Most students prefer to investigate science problems for which they do not know the answer/s.

Aspects of the Cognitive Apprenticeship Model of Instruction

Teacher Modelled Investigations

The students were not enthusiastic about the teacher modelled investigations. For the teacher modelled pitch of a closed pipe investigation each group commented on their lack of learning and implied that this was because they were not directly involved. These data support Assertion 28: Students perceive that they learn less about doing investigations from teacher modelled investigations than from student investigations. In addition they support Assertion 29: Students perceive that teacher modelled investigations contribute less than student investigations to their learning about *Planning investigations, Conducting investigations, Processing data* and *Evaluating investigations.* With regard to the teacher modelled investigations Group TN said, "We didn't like it. We didn't do anything. It was pretty boring. We didn't learn anything. We learn more if we do it ourselves." These sentiments were also expressed by Group TC who said, "We didn't learn much. We didn't do it," and Group SC who said, "We didn't actually do anything. She (Mrs Cross) did it on the board." Group SC said that the investigation was easier than other investigations and implied that it was because they already understood the prior science content knowledge. "It was easy because it's common knowledge with sounds" they said, referring to the relationship between length of the closed pipe and the frequency of the note. The idea that students' familiarity with the science content knowledge

affects the difficulty of the investigation is expressed as Assertion 52. This supports claims by other researchers that prior knowledge affects students' success in preforming investigations (Germann & Aram, 1996; Gott & Duggan, 1995). In addition, Group SC acknowledged that, "We learnt that the planning is difficult. Doing it in your head is one thing but writing it down so that someone else can do it is totally different." This suggests support for Assertion 37 which states that students' plans lack sufficient detail to adequately describe the procedures.

Assertion 52: The difficulty of investigations is influenced by students' prior conceptual knowledge associated with the investigations.

For the teacher modelled acid and carbonate investigation Groups TN and TC referred to becoming aware of competencies such as *Being organised* anc *Managing time*.

Group TN: Mrs Cross showed us that we need to set up the equipment really well at the start. To get organised at the start. We revised hypothesis writing but she did not have enough time to do trials.

Group TC found it difficult to remember the investigation and added,

> Did we do that? (There was a lot of hesitation.) I don't remember anything. (Gradually they recalled the investigation.) You have to get organised before you start. You have to be quick with the stopper. (Put the stopper in the conical flask so that the carbon dioxide does not escape.) The limewater goes milky.

Group SC agreed that they didn't learn much from the teacher modelling and they said that they learned more when they performed the investigations. This adds support to Assertions 24, 28 and 29. They also stated that they learned more from new tasks (Assertion 19).

Group SC: I do (learn) when it's new. I learn when I watch but it really sticks when I do it myself. It solidifies in my mind if I do it myself. (Jody)

With further questions they indicated that they learned about *Controlling variables* and *Managing time*.

Group SC: Keep everything constant and have the same person testing each thing. She (Mrs Cross) kept on stressing this. We waited for the first milkiness in the limewater. Mrs Cross was the person watching. She said to stop timing. Someone else was timing. She told us that we had to conduct more than one trial. W . looked at the more concentrated the acid the faster the calcium carbonate reacts. We didn't get many results down. We got enough to prove the hypothesis right. We got the graph but we didn't quite finish the rest of it. You have to be quick and get onto it. Obviously she didn't have enough time to finish it. So if we took as long as she did then we would take too long."

These data support questionnaire data (Table 7) in which students rated the teacher modelled acid and carbonate investigation ahead of the teacher modelled pitch of a closed pipe investigation in terms of contributing to learning science content knowledge and learning about doing science investigations. Overall, the lack of enthusiasm for and learning from teacher modelled investigations, particularly the closed pipe investigation, may be associated with teachers' perceived difficulties in implementing the modelled investigations, Assertion 40. Suggestions to improve the implementation of the teacher modelled investigations in this study were made in Chapter 6 and are summarised in Assertions 41: Teacher modelling of investigation competencies may best be implemented with small groups of students so that teacher guidance can be (a) provided when needed by the group, (b) highly focused on the difficulties encountered by the group, and (c) at a personal level. Assertion 42 also makes suggestions about teacher modelled investigations: The focus of teacher modelling should be on developing a small number of competencies. Support for implementing the modelled investigations differently is evident from the comments of Group TN. "Just do it (the investigation) and ask the teacher when you need help. Let them (the students) do it and (have the teacher) say that you're doing the wrong graph. We learn from doing the investigation and we learn from our mistakes."

Coaching and Articulating

Eight of the nine group members were positive about learning from discussions with their peers. Groups TN and SC said that their peers used more appropriate language in their explanations than their teachers. From Group TN Gemma said, "They (students) help you with everything," and implied that the teachers' help was reserved because they "only point you in the right direction." These data support questionnaire data, Assertion 25: Most students perceive that they learn more about doing investigations from talking with their peers than from talking with their teacher. In contrast, Olive stated that she preferred to do things by herself rather than to discuss matters with the group because she found it hard to reach consensus with the group. No new assertions were formulated from these data.

Group TN: They (students) realise what is going wrong and they use your kind of words to explain things.
I like talking to my friends but I don't like talking to the teacher (Gay).

Group SC: If I don't know how to do something the next person can usually help. The whole group really has to do it on their own because the teacher can't really go around helping everyone at once. It's really like doing it on your own. After the first one (investigation) we'd got the format. There were things that were always the same that you could go on with. Sometimes you don't understand what the teacher says. Sometimes it helps if another student explains because their language is better than the teacher's.

Group TC: We learn about the same from teacher and the students. It's easier to learn if you **do** things yourself than if you talk with friends and with the teacher. The teacher knows what we should do. They can put you on the right track (November)
Other people never agree with what you think. Therefore, I prefer to do things by myself. It's less frustrating. The teachers help you but they don't do it for you. (Olive)

In terms of the cognitive apprenticeship model it is interesting that students did not acknowledge their teachers to be expert coaches. This may be a consequence of a mismatch between teachers' and students' views about the amount of guidance that should be given to students performing investigations. For example, in Chapter 6, Mr Brogo thought students should work things out for

themselves and Mrs Grant preferred to help students. A range of teacher opinions about helping students is acknowledged in Assertion 44: Teachers have different views on how much guidance to give students performing investigations. It may be that "discovery learning" in which students "are invited to 'explore' and to 'find out what you can', (Hodson, 1996, p. 116), although fashionable in the 1970s still influences teachers' perceptions of their role in investigative science classes. Also, it may be that teachers are too busy with classroom management to effectively coach students. The effectiveness of formative assessment during the investigations has previously been questioned in Chapter 6 because without the formal assessment of students IPRS teachers perceived that they lacked feedback on students' performances.

Scaffolding and Fading

The groups were in agreement that the scaffolded Investigation Planning and Report Sheets were helpful in their learning. The scaffolded IPRS (Appendix F) was used for all investigations except the pretest and the posttest investigations when the IPRS with less scaffolding was used (Appendix D).

Some students suggested improvements that could be made to the scaffolded IPRSs such as rewording the planning section to reduce the repetition of having to briefly outline a plan and later describe the plan in detail. These comments are consistent with those of Mrs Grant and Mrs Cross in Chapter 6 (p. 168).

The following comments reveal that students found the scaffolding helpful in their learning, and Groups TN and TC recognised that the scaffolding could be withdrawn as the program progressed and their competencies improved, Assertion 53.

Group TC: It was a bit repetitive. eg the planning section. (August)

Group TN: It was easier when there were lots of questions to fill in. They were helpful in showing us what to do. We had no idea what we were doing in the lipase investigation. We can use the other sheets (IPRSs) now. (November)

Group SC: The worksheets (scaffolded IPRSs) are good. They help you to set things out. (November)

Group TC: The ones (scaffolded IPRSs) at the beginning were good because they were so detailed. They were teaching us what to do. It gets a bit repetitive once you get the hang of it. It takes a long time to fill in. It's good to get the other ones (with less scaffolding). But we need more detail at the beginning because we didn't know what we were doing. It's hard to find the variables at the beginning when you don't know much about the experiment. We hated writing up the plan. At the beginning you need lots of questions, lots of detail. At the end you can do it all by yourself. (November)

Assertion 53: Students perceive that scaffolded Investigation Planning and Report Sheets as implemented in this study are helpful in learning to perform investigations, and that less scaffolding is appropriate at the end of the program.

Self-Reflective and Metacognitive Skills

The worksheet on marking an investigation, the questionnaire that was completed after each segment of 10 lessons, and the *Evaluating the Findings* section of the IPRS required students to exercise self-reflective and metacognitive skills. Students in Class SC were involved in the assessment of their investigations and this involved self-reflection also, and is discussed in the following section of the Chapter

The students were asked about their perceptions of the worksheet on marking an investigation which required them to assess an investigation completed by a previous Year 9 group of students (Appendix G). Group TN responded negatively to the task. Group SC was more positive and indicated that they were able to reflect

on their performances in relation to others. (Group TC did not do this task.) No assertions were formulated from these data.

Group TN: Didn't learn anything. We wouldn't have done all the graphs because we couldn't be bothered. We didn't understand what was happening. (August)

Group SC: We learnt how you do a marking scheme and what teachers look for. In our next one we thought about it (the worksheet on marking an investigation) to see if we had done what the teacher said we should do when she went over it on the board. You can learn from people's mistakes. If you put us up against these people we would do much better than them. They didn't plan. The variables aren't the same (controlled). They did the hypothesis and the conclusions were different. They did it wrong. What they found just supported it (their hypothesis). (August)
I hadn't thought about how to mark before. I didn't think it was a waste of time. (Harue)

Each group of students said that they didn't like the questionnaires (Appendix M). They indicated that they did not like documenting what they had learned and this is illustrated in their comments.

Group TN: It's a waste of time really. We don't like it. There's no point to writing down what you learn. You just learn it.

Group TC: We don't like the questionnaire. It is silly, pointless and doesn't help. It doesn't do anything for us. We don't like giving reasons. It's stupid and aggravates us. (August)

Group SC: I hate it when we come to the end and we do the questionnaire. I can't remember what I have learnt from each experiment. I don't really know. I know what I knew at the beginning and I know what I do know now. I don't really know what to say. I thought it was boring. I think it's bad. It didn't help me at all. (Harue, August)

A few students stated that they did not like the last page of the IPRS (Evaluating the findings) because they didn't know what to say and that they found it pointless. This supports teachers' perceptions that Year 9 students are not sufficiently mature to reflect on their learning, Assertion 45. In addition, this view is in accord with questionnaire data (Assertion 21) that indicates that "Students do not identify *Evaluating investigations* as perceived learning about investigation competencies."

The fact students did not like answering the evaluative section of the IPRS may also indicate that they needed additional instruction on how to answer the questions and the possible benefits of questions that direct them to think about their learning.

Group TC: We don't like the last section. We hated the end questions. The way science investigations are carried out. I don't care. The purpose of the investigations is to learn the steps and group work and stuff. We didn't do the last few pages. We can't think of what to say. It's pointless because you don't want to do it you say you can't think of anything. (August)

The small sample of Year 9 students interviewed about the worksheet, the questionnaire, and the *Evaluating the findings* section of the IPRS, conveyed that they did not like participating in self-reflective and metacognitive practices. Students' views about self-reflection and metacognition are complemented by the teachers' perception that students are insufficiently mature to reflect on their learning (Assertion 45). The students' lack of recognition that these practices may contribute to learning are expressed as Assertion 54.

Assertion 54: Most students do not recognise that self-reflective and metacognitive tasks could be of benefit to their learning.

Assessment

Teacher Assessed Norm Referenced Assessment: Group TN

Group TN had their work marked by two teachers (Mrs Cross during May and November, and Mrs Grant during August) and were assigned a grade based on the distribution of the class results. They said that during August they did not learn much from making corrections, and during May and November they were more

positive. On these occasions they found it helpful when the teacher went over the investigation in a whole-class setting. The value of whole-class feedback has been expressed as Assertion 55. From the teachers' perspective, Miss Mills (p. 172) commented that she found this an efficient way to provide feedback although she questioned the students' interest because the investigation had been completed.

May: It was helpful. We found out how to improve when she went over it in class.

August: We should do more trials. We got a B+.
No! Not much. (This was their response to learning from making corrections.)

November: The marking was quite helpful. We got a grade and we also got a comment because she wrote what we did wrong. We learnt to answer the back page (of the IPRS). She said that we need to be more specific and she wrote the question number for what we did wrong. It was helpful when we went over it in class. We spent about half a lesson on corrections. It (making improvements) was OK going back and writing things. I think we knew how to do it better.

Assertion 55: Some students believe whole-class feedback on investigation performance is helpful in improving performance.

Teacher Assessed Criterion Referenced Assessment: Group TC

Group TC had their work marked by Miss Mills during May and November and Mr Brogo during August. The teachers use the criterion referenced marking sheet (Appendix I) and assigned a grade of A, B or C for each criterion. They also wrote comments for some criteria. When specifically asked by the researcher about the criteria used for their assessment the group revealed that they found the criteria useful in providing feedback on their progress and this has been expressed in Assertion 56.

May: It's good! It shows us what we need to improve on. We didn't make any corrections. I can't remember going over it in class (Jessica).

August: We went better on the ukulele investigation. We got only one C. For the electromagnet we got five Cs. We couldn't understand the teachers' writing so it was hard to make corrections. We knew which areas we should improve.

November: It was useful. The teacher wrote a comment next to it (the grades) so that we could go back and look at it the next time. It was helpful because we know what to improve on next time. We know exactly what we had done wrong.

Assertion 56: Some students believe the teacher assessed criterion referenced assessment as implemented in this study is helpful because it highlights investigation competencies that could be improved.

Assertion 56 was generated in response to a specific question asking students what they thought of the criterion referenced assessment used to assess their investigations. The assertion contradicts Assertion 33: Students do not acknowledge that criterion referenced assessment is helpful in providing feedback on investigation performance. Assertion 33 was generated from questionnaire data asking how much the student had learned from the assessment and, in particular the follow-up open-ended question asking students to give examples of what they had learned. No students in Class TC or Class SC acknowledged the criteria. Previously (p. 146) it was documented that this was surprising given claims (Biggs & Moore, 1993; Gipps, 1994; Harris & Bell, 1994; Popham, 1992) for numerous strengths of criterion referenced assessment, however, it needs to be noted that there was no direct question about criterion referenced assessment in the questionnaire.

Student Assessed Criterion Referenced Assessment: Group SC

Group SC marked their own investigations against a master answer sheet. They used the same criteria as students in Class TC to determine grades of A, B or C for each criterion. During May and November they expressed uncertainty about

judging their performance and this is expressed as Assertion 57. Harue and Jody commented that they would have liked the teacher to confirm their assessment. They tended to base their grades on the degree of difficulty they encountered, assigning themselves an A if they perceived that they had no problems, and a C if they had 'trouble' with some part of the investigation.

May: We did meet the criteria. We could have improved at developing an overall plan and being more organised. (They discussed what information they used to judge the criteria.) Sometimes we were not sure if it should be an A or B. We gave ourselves a B for *Data is accurate* because it depended on how we positioned the ukulele and the different ways we strummed it and held it. We had to be careful because otherwise we got voices in the recording. We had to be careful capturing the results.

August: We got As because we got this one right (Panadol). We got Bs and Cs in this (catalyst investigation) because we had trouble collecting the data. We got As for things that didn't have anything to do with the results.

November:

Harue: I don't know if we liked it (doing our own assessment). I would have liked the teacher to do it as well so I could see what/how I should have marked it because it was kind of weird marking your own thing because I didn't know if it was right or wrong. Getting the sheets helped with the marking (because the correct responses were written out).

Christie: You had to go right through it and then go and do the A, B and C stuff. The teacher would be more objective. We got harder on ourselves as we went on through the year. I think we should do some marking but not every time. We don't know if it's right or wrong.

Jody: Probably I learnt from that marking I think ... But with the teacher you can go and ask why you got this mark or that mark. I would like to have the teacher confirm that I am doing it right.

Assertion 57: Students are uncertain about judging their investigation performance based on predetermined criteria.

The following dialogue illustrates how students obtained feedback from the master answer sheet. Their comments clearly indicate that they reduced the cognitive demand of correcting their work to a minimal level. Mostly they graded

the criteria and added information from the master answer sheet without thought of improvement. As a result, potential benefits from feedback opportunities were minimal. Therefore, without teacher guidance it appeared that students who assess their own science investigations do so in a superficial way and have reduced opportunities for teacher feedback on their performance (Assertion 58). The lack of meaningful clarification of students' errors from this type of assessment was raised by teachers (Assertion 50). They stated that a major concern with student self-assessment was that it was difficult to obtain information about the students' performances and hence difficult to give them appropriate advice on how to improve.

Harue: "We are kind to ourselves. (They laugh.) We didn't have to word it (the feedback) to ourselves. We just filled in the boxes.

Jody: "We usually got most of it right. A few times we added more in. We didn't really think about it. We just copied it in."

Assertion 58: Students' self-assessments of their investigations are superficial and they offer reduced opportunities for teacher feedback.

Summary of the Chapter

Three groups of students were interviewed after each 10 lesson segment. Eight assertions were proposed and these are summarised in Figure 20. Two assertions were formulated about students' perceptions of investigation competencies. The notion that students prefer to perform investigations that lead them to the construction of scientific conceptual knowledge that is new to them, affected their perceptions of the usefulness and purposefulness of investigating.

Based on the group interview data, assertions were made about scaffolding and fading (Assertion 53), and self-reflective and metacognitive skills (Assertion 54).

Four assertions were made about the assessment procedures. When students assessed their own investigations they were uncertain about assigning a grade for their performances, they were superficial in their approach and they had reduced opportunities for teacher feedback on their performance.

Assertion 56 contradicted the previously generated assertion, Assertion 33: Students do not acknowledge that criterion referenced assessment is helpful in providing feedback on investigation performance.

The student interview data has provided supporting evidence for assertions formulated from other data sources. The following assertions about students' investigation competencies, grounded on the IPRS data were supported, 1, 2, 4, 7, 8, 12 and 14. From the student questionnaires support for Assertions 19, 21, 22, 24, 25, 28 and 29 was presented, and from the teacher interviews there was support for Assertions 36, 37, 38, 40, 41, 42, 45 and 50.

Assertions

Investigation competencies

51 Most students prefer to investigate science problems for which they do not know the answer/s.

52 The difficulty of investigations is influenced by students' prior conceptual knowledge associated with the investigations.

Aspects of the cognitive apprenticeship model of instruction

(a) Scaffolding and fading
53 Students perceive that scaffolded Investigation Planning and Report Sheets as implemented in this study are helpful in learning to perform investigations, and that less scaffolding is appropriate at the end of the program.

(b) Self-reflective and metacognitive skills
54 Most students do not recognise that self-reflective and metacognitive tasks could be of benefit to their learning.

Assessment

(a) General
55 Some students believe whole-class feedback on investigation performance is helpful in improving performance.

(b) Teacher assessed criterion referenced
56 Some students believe the teacher assessed criterion referenced assessment as implemented in this study is helpful because it highlights investigation competencies that could be improved.

(c) Student assessed criterion referenced

57 Students are uncertain about judging their investigation performance based on predetermined criteria.

58 Students' self-assessments of their investigations are superficial and they offer reduced opportunities for teacher feedback.

Figure 20. Summary of assertions developed from group interviews with students

CHAPTER 8

AUDIO AND VIDEO DATA

Overview of the Chapter

This Chapter describes audio and video data that were collected from the three classes involved in the research and makes assertions about the cognitive apprenticeship model of instruction, and the different assessment regimes implemented in the study. Supporting data were found for previously generated assertions and no disconfirming datum was found. The data obtained from the audio and video tapes provide rich, authentic descriptions of the classroom learning milieu and the ways groups of students performed science investigations. Hence, they provide an additional dimension to the study. Teacher instruction in a whole-class setting and the work of Groups TN (Kim, Tammy and Gay), TC (Jessica, Olive and Gemma), and SC (Christie, Jody and Harue) from the Classes TN, TC and SC respectively were recorded on audio and video tapes. Written work from the groups (worksheets and Investigation Planning and Report Sheets) was examined in conjunction with the audio and video tapes.

Aspects of the cognitive apprenticeship model of instruction are addressed under the headings teacher modelled investigations, scaffolding and fading, teacher guidance, and articulating. Although discussions have been shaped under these headings, the components of the cognitive apprenticeship model are not mutually exclusive. Teacher guidance, for example, is associated with articulating and scaffolding. Self-reflective and metacognitive skills were not evident from the audio and video data. Finally, the Chapter attends to the impact of the assessment regimes; teacher assessed norm referenced, teacher assessed criterion referenced and student assessed criterion referenced assessment. Throughout the Chapter, scenarios

extracted from the audio and video tapes are presented. They include direct transcriptions of dialogue and/or descriptions of the activities in which the teachers and students were involved.

Aspects of the Cognitive Apprenticeship Model of Instruction

Teacher Modelled Investigations

Four scenarios are presented to describe how the teacher modelled investigations were implemented by Mrs Cross, Miss Mills, Mrs Grant and Mr Brogo. Based on these data five assertions were formulated.

Scenario 1.

Mrs Cross modelled the acid and carbonate investigation in a similar way on two occasions, during May to Class TN and during August to Class SC. She devoted two lessons to the modelled investigation. She maintained an efficient, brisk, business like approach that was highly focused on students completing the Investigation Planning and Report Sheet (IPRS). She made it clearly understood that she expected students to write answers for each section of the IPRS, and then she asked some students to read out their responses. In both classes there was a high level of task engagement and this was probably because the students were required to complete the IPRS. Mrs Cross took care to ensure that students recorded the correct answers and at times she dictated information such as the hypothesis, and at other times she wrote the information on the white board, such as the method and the diagram. Occasionally she highlighted difficult parts of the investigation. For example, she said measuring the dependent variable was quite hard and they must think about what to measure. She added that if they were measuring "rate" then they were looking at "time" and that in this instance they were measuring the rate of production of carbon dioxide gas. She did not discuss other ways the rate could be measured, such as the change in pH of the solution or the change in mass of the calcium carbonate. She briefly discussed the importance of trialing and that students should use trials to improve their data collection and to work out what they were

doing. She mentioned that scientists repeat trials hundreds of times. Mrs Cross stressed the need to be organised, to have everything set up at the outset, and the need for one person to do a specific aspect of the data collection. She also focused on the need to control variables to ensure a fair test; the need to repeat tests to improve accuracy; and the role that preliminary trials play in deciding how and what to measure.

For both classes, groups of students demonstrated the trials to the class. Mrs Cross did the timing and organised the set-up of each trial. This section of the lesson took a long time and was quite difficult to manage because she was trying to collect the data, and to keep the class focused on writing responses on the IPRS and drawing their graphs. When the noise level gradually crept up, Mrs Cross called the students back to order by asking them to write up another section of the IPRS. During May she collected only one set of results and the next lesson she presented them with additional data that they were asked to graph for homework. During August she managed to complete both sets of trials, however, because the class was short of time the students were asked to finish off the investigation at home.

Scenario 2.

Miss Mills devoted two half lessons and one full lesson to the teacher modelled acid and carbonate investigation. For the first half lesson she organised students around the teacher's bench and conducted a 35 minute long question and answer session during which she revised the equation for the acid and carbonate reaction and the test for carbon dioxide. The students appeared to remember very little and she gradually elicited information from them. In addition, she discussed the variables that needed to be controlled. The next day she again organised the students around the front bench and then conducted four trials. It was difficult for the students to observe the reactions in the test tubes because they were too far away. Miss Mills conducted two trials with 0.2 mol L^{-1} HCl. The first trial with granules of calcium carbonate did not produce sufficient carbon dioxide and neither did the second with powdered calcium carbonate. She mentioned shaking the test tubes and explained that this would need to be done for all the tests. She then said that she would use the 2 mol L^{-1} HCl with calcium carbonate granules. She demonstrated this to the class and explained that she was still conducting the preliminary trials. She also stressed the need to repeat trials to make sure the results were accurate, and the difficulty in judging when to measure the dependent variable. During this time the class became quite restless. Some students kept touching other students and then pretended that it was not them. Miss Mills sensed their inattentiveness and

continually told them to pay attention. For the final 10 minutes of the lesson the students moved back to their seats and were instructed to fill in the IPRS up to question 3. Miss Mills stood at the front of the class and, because the students were unable to remember the details, they continually raised their hands to ask questions as they wrote their responses. They asked a lot of procedural questions such as how much acid and limewater was used. The following day, in a whole-class setting, Miss Mills discussed some students' responses to the first sections of the IPRS. She also provided them with a diagram of the experiment to copy from the white board. The students were unable to answer the questions about processing data and evaluating the findings because insufficient data had been collected.

Scenario 3.

Mrs Grant performed the teacher modelled pitch of a closed pipe investigation during May to Class SC, and during August to Class TC. At the beginning she explained that as a result of the modelling she hoped they would "pick up some tips about the way to do investigations." Her time management of the modelled investigations differed markedly from May to August. During May she devoted two lessons entirely to the modelled investigation, however during August she devoted only one lesson because she spent the first lesson revising the terminology associated with investigations, including hypothesis, independent variable, dependent variable and controlled variables. She may have perceived that the revision of the terminology would be more beneficial because it had been three months since the students' last investigation. The modelled investigation during August was very rushed compared with the May investigation although the same investigation procedures were addressed.

Mrs Grant used the IPRS to provide structure for the lessons. Typically, she used question and answer lesson segments to promote discussion. For the data collection, students gathered around the computer and tried to catch a glimpse of the computer screen, however, it was difficult for them to see. The students then recorded relevant information on their IPRS. This approach was different from that used by Mrs Cross because Mrs Cross first asked the students to fill in their own responses, and then followed-up with a discussion in which the 'correct' responses to the IPRS were given.

On both occasions Mrs Grant addressed a number of data gathering issues. These included

(a) how to use the computer to collect data,

(b) the number of trials that should be conducted. She said that a golden rule was to, "Do as many trials as you need to do to get reliable or certain results This will vary with every experiment that you do."

(c) the need to control variables by maintaining a constant "blow" over the top of the pipe. (If students blew too hard then the second harmonic was produced and the frequency was double the fundamental.)

(d) different ways to measure the length of the pipe,

(e) the number of different lengths of pipe that should be tested to establish a pattern,

(f) the length of time over which the computer sampled the wave,

(g) the need to identify and attempt to explain spurious data, and

(h) the omission of data that was known to be wrong such as errors caused by technical problems.

Numerous data were collected, however, Mrs Grant placed a previously constructed data table and graph on the overhead projector for the students to examine. At this stage the lesson concluded and she asked the students to finish up to question 5a for homework. At the start of the following lesson she directed students to complete question (6 d ii) on "the way science investigations are carried out." In response to this question she said they should do "as many trials as were needed to achieve a pattern in their results" and that they "should not ignore data unless it was an impossible result that was due to a technical error" and that they "knew the data to be wrong."

Scenario 4.

Mr Brogo modelled the pitch of a closed pipe investigation to Class TC during August. During the first lesson he talked about variables that would affect the pitch of the note, the need to change only one variable for the investigation, the need to average results to improve reliability and the fact that "we make judgements using our senses." He also explained that in the past scientist had to make their own equipment before they could conduct experiments. The students did not complete any specific tasks during the lesson and appeared to lose interest. The question and answer session continued for the whole lesson. During the second lesson, after reviewing the preceding lesson, Mr Brogo asked the students to write responses on their IPRSs. About half-way through the lesson he showed them how to collect the data from the computer. He repeated the trials three times and as the lesson drew to a close he said that he would continue to collect data after the lesson and give the students a copy of the data the next day. The students were instructed to finish off as much as they could for homework. During a subsequent lesson he provided them

with additional data and later they submitted their IPRSs for comment. Mr Brogo responded by providing very detailed written feedback.

In summary, the audio and video data illustrated that teachers attempted to model the whole investigation and previous difficulties that were identified from their interviews (p. 161-165) were confirmed. There was limited modelling of teacher thought processes, however, there was considerable guidance about what to do and how to proceed. It was confirmed that teacher modelling was time consuming, students became off-task, not all students could observe the data gathering, and that the setting up of the equipment took a long time (Assertion 40). The scenarios also indicate that during the modelled investigations, maintaining student engagement for a protracted period of time was difficult for both teachers and students. Also, at times the equipment was difficult to manage, and it was difficult for students to observe at a distance. Therefore, the data support previous assertions that teacher modelling may best be implemented with a small group of students, provided when needed and focused on the needs of the group, at a personal level (Assertion 41), and limited to the development of a small number of investigation competencies (Assertion 42).

Students' attentiveness during extensive question and answer lesson segments was observed to decline, particularly during lessons when approximately 35 minutes was devoted to this teaching strategy (Scenarios 2 and 4). Student attentiveness and on-task behaviours were more evident in lessons where the IPRS was used to structure the lessons (Scenarios 1 and 3) and where students wrote their **own** responses (Scenario 1) on the scaffolded IPRS prior to class discussions. Feedback was then provided by the teacher after a number of students were invited to read out their responses to the class. The notion of using the IPRS to improve on-task behaviour is expressed as Assertion 59.

Assertion 59: During teacher modelled investigations on-task student behaviour is maintained by requiring students to complete Investigation Planning and Report Sheets.

The information that was modelled to the classes differed, depending on the teacher and the nature of the investigation. For the chemistry investigations described in Scenario 1, there was an emphasis on the need to repeat trials, control variables by having the same person perform a specific task, and to be organised. In Scenario 2, which was also a chemistry investigation, preliminary trialing and the variables that needed to be controlled were discussed. During these investigations the organisation of equipment was an important feature. In contrast, there was more potential to address broader aspects of data gathering in the teacher modelled physics investigation as exemplified in Scenario 3. The idea that different investigative tasks provide opportunities for teachers to model different investigation competencies is expressed in Assertion 60.

Assertion 60: Different investigations provide opportunities for teachers to model different investigation competencies.

Mrs Grant used second hand data to save time and to effectively model data processing techniques (Assertion 61). She presented tables and graphs that she had previously constructed on the overhead projector to effectively improve the pace of the lesson and maintain student engagement. In addition, the use of the overhead projector enabled her to supervise students as they attended to their tables and graphs.

Assertion 61: Second hand data can be used effectively to model data processing competencies.

Whole-class management appeared difficult during the data collection phase. The difficulties were exacerbated when students moved from their seats to have a better view of the investigation (Scenario 2). This provided them with opportunities to touch and nudge each other. Also, management difficulties were more evident when the data collection extended for a long period of time and repeat trials were conducted (Scenario 1). During the data collection the teachers had to focus on the management of the equipment as well as the management of the class. These ideas are expressed in Assertions 62 and 63.

Assertion 62: Moving students from their seats so that the data gathering can be observed may result in increased off-task behaviour.

Assertion 63: Students appear to lose interest when watching repeated or replicated trials.

Scaffolding and Fading

During the instructional program scaffolding was provided for students by the Investigation Planning and Report Sheets and fading was achieved by reducing the amount of scaffolding for the final investigation. The audio and video data indicated that the IPRSs were also used by teachers to organise instruction and to provide a focus for student learning and this is expressed as Assertion 64. This was particularly evident during the teacher modelled investigations in Scenarios 1 and 3.

Assertion 64: Teachers use the scaffolding provided by Investigation Planning and Report Sheets to organise instruction and to provide a focus for student learning.

In addition, the audio and video tapes indicated that students used the IPRSs as checklists to ensure that they had completed the requirements for the investigations. Groups TN, TC and SC were observed checking that they had responded to the sections on the IPRS and this is expressed as Assertion 65. This assertion supports the student interview data in which students perceived that the scaffolded IPRSs were helpful when they were learning how to perform investigations (Assertion 53).

Assertion 65: Students use the scaffolding provided by Investigation Planning and Report Sheets to ensure that they complete the requirements for investigations.

Teacher Guidance

Examples of teacher guidance are portrayed in Scenarios 5 and 6. They illustrate that teachers differ in the type and amount of guidance they give to students. Mrs Cross (Scenario 5) tended to give specific information and direct replies to questions as she told students what to do and how it was to be done. An example of this occurred when she told students to measure the time taken for photographic film to become clear in the trypsin reaction. She did not tell the students everything. She did not provide so much information as to remove the problem solving component from the investigation, however, she was not observed trying to lead students to the answer by a series of questions. This is consistent with the way that Mrs Cross conducted the teacher modelled investigations.

In contrast Miss Mills' approach was more consistent with Socratic dialogue described by Brown and Palinscar (1989) and the concept of scaffolding espoused

by Collins et al. (1989). In Scenario 6, Miss Mills used Socratic questioning to elicit a correct response for the trypsin investigation. This form of guidance also is similar to that used in the study by Roth and Bowen (1995) in which guidance for students in open laboratory tasks was in the form of carefully constructed questions to provoke problem solving.

Scenario 5.

Mrs Cross gave clear, precise and brisk initial instructions. Then she advised, "Work out the problem. Work out what your group will investigate. Do even one trial. You need to see how it (the photographic film) goes clear and how long it takes to go clear. Get a rough idea of what you are timing and what you are going to investigate." She then moved around the class talking to **each** group to check that they could get started.

During the investigation Kim said, "What if it goes clear before two minutes? Do we stop it? "Yes," said Mrs Cross without further explanation. Later Mrs Cross looked at their results. "What does B mean?" she said as she looked at their table of results. "It's the boiling water," replied Tammy. Mrs Cross replied, "You have to be more accurate than that. You have to record the temperature." Kim and Gay appear embarrassed by their omission and said the investigation was stupid.

Scenario 6.

Miss Mills instructed Class TC about the trypsin investigation. She started by telling them that at the beginning of the year they looked at lipase breaking down fats in milk and that now they will look at trypsin breaking down protein. Initially she did not tell students that the purpose of the photographic film was to provide a source of protein. She asked them to read the background information and to find this out. She then asked for students to respond and developed their responses further by asking, "Why do they have the sentence that says to do the investigation at 37 °C?" She requested that the students "think carefully" and allocated an appropriate wait time before a student responded.

These scenarios suggest that teachers differ in the ways they help students to perform science investigations and this is expressed as Assertion 66. This assertion

complements and supports previous assertions that teachers have different views on how much guidance to give students performing investigations (Assertion 44); and that some students prefer teachers who give full explanations and some prefer teachers who encourage and assist them to solve their own problems.

Assertion 66: Teachers differ in the type and amount of guidance they give students performing investigations.

Scenario 7 is an example of teacher guidance which had limited success in terms of student learning. The scenario portrays Group TC's interactions during the Panadol investigation in which they investigated the time taken for the Panadol to dissolve in different solvents. Miss Mills resolved the Group's immediate difficulty in determining whether to plot a line graph or a bar graph, and they successfully plotted a bar graph. An examination of their IPRSs showed that they succeeded with this graph. Subsequently with the trypsin investigation, they decided that a line graph of temperature against time was needed but they were unable to successfully draw the graph because they could not determine the scale for the axis. Scenario 7 highlights difficulties associated with graphing which have been previously reported by Roth and McGinn (1997). As well, it illustrates that the teacher did not foresee problems that students would encounter constructing a line graph for the trypsin investigation. Hence, the scenario also highlights a major difficulty for teachers in accessing students' beliefs and understandings.

Scenario 7.

Jessica: So it's a bar graph?

Gay: No!

Jessica: How do you know then?

Gay: Because Pam told me. I think we're right. If it's a bar then they're changing and if it's a line they're constant.

Jessica: I would have thought it would be a bar, but don't worry then

Gay: You do a bar then

Jessica then asks Miss Mills and she goes to their bench and she says to the group, "Well ask yourself if the data is continuous." Jessica guesses "Yes!" Miss Mills then replies, "What's the answer then?" Jessica then acknowledges that she does not know what is meant by continuous data.

Jessica: Continuous? What do you mean continuous?

Gay: Constant like.

Jessica: Yes! Does that mean there's gaps in it? Oh! Does that mean that it goes up in a straight line.

Miss Mills then used a scaffolding of questions to lead the students to distinguish between continuous and discrete data. She gave examples of continuous data and discrete data and emphasised that with continuous data it would be possible to take measurements between the data points. Jessica enthusiastically says, "I get it now," and Miss Mills leaves. Jessica and Gay then start to sing a song. Later they successfully plotted a bar graph for their data (time taken for the different solvents to dissolve Panadol).

Subsequently, as indicated on their IPRSs for the trypsin investigation, errors emerged with their line graph. They divided the x axis into three equal segments and labelled the centre of the segments 30, 37 and 45 °C respectively, as if they were dealing with discrete data. The coordinates for temperature and time taken for the reaction were plotted at the **end** of the temperature segment and connected with a line. It appeared that although they realised temperature was continuous data they did not know how to construct a scale for the temperature data. Therefore although Miss Mills helped them with their immediate problem of deciding which type of graph to plot, she did not anticipate that they could not construct a scale for the line graph.

This scenario highlights a major difficulty for teachers in accessing students beliefs and understandings. Previously researchers have documented that teachers can obtain feedback from students through observations (Collins et al. 1989) and questions (Gipps, 1994; Javela, 1996; Roth, 1995; Torrance, 1993). Notwithstanding this feedback to the teacher, teacher judgements based on these data sources may be "incomplete, fuzzy, qualitative and based on a limited range of

potential criteria" (Gipps, 1994, p. 130). For teachers to access students' understandings it appears that they need to be aware of potential difficulties that students may experience and use these difficulties as a framework for inquiries about students' progress and subsequent coaching decisions. When coaching, teachers do not always gather appropriate information about students' understandings to form a basis for suitable coaching or feedback to the students on their performance. This notion is expressed as Assertion 67. Even when teachers construct careful explanations, students form unintended meanings that are not consistent with the teacher's explanations in ways that are outlined by Roth and Bowen (1995).

It may be argued that coaching would be better tailored to students' needs if students were able to identify deficiencies in their understandings and ask teachers questions so that appropriate feedback could be provided. For students to be able to ask such questions it is postulated that well developed metacognitive skills would be required and that students would need much greater access to the teacher to have their questions answered.

Assertion 67: Teachers do not always gather appropriate information about students' understandings to make informed decisions about ways to coach students.

Articulating

The cognitive apprenticeship model of instruction purports that learning is facilitated through articulation between an expert (teacher) and novice (student). This section, however, focuses on student-student interactions. Therefore it addresses learning from a collaborative group work perspective rather than an expert-novice perspective. Scenarios 8, 9 and 10 describe the way the groups completed their work and portray typical interactions between the members of

Groups TN, TC and SC respectively. Following the scenarios emerging themes regarding articulation are addressed and three assertions are presented.

Scenario 8.

Group TN comprised Kim, Tammy and Gay. Tammy was the motivator and the group organiser. Kim contributed sound ideas and Gay was quite negative and did as little as possible. Gay spent a lot of time adding graffiti to her worksheets and files, and during one lesson she did this for 25 minutes. Typically, Tammy and Kim arrived at the bench and collected the equipment straight away. Gay was always late because she engaged in social discourse with other groups before arriving at the bench. For the posttest investigation she arrived 13 minutes after Kim and Tammy. Notwithstanding this they appeared to be the most efficient and capable of the three groups studied. Kim and Tammy got things done because they worked cooperatively. Tammy was most often the group leader with Kim willing to follow and contribute. Gay, on the other hand, complained about everything and was quick to grab the other girls' worksheets and IPRSs so that she could copy their work. They were not overtly off-task because they worked comfortably within the bounds of classroom behaviour. Their discussions about the investigations were about the procedures required to complete the task at hand. They didn't discuss issues of accuracy or ways to improve the investigation. They were not observed engaging in discussions that could be deemed capable of contributing to the development of procedural understandings. They seemed to make statements rather than discuss or resolve issues. An example of this follows.

Kim: Do the 1 mL (measurements) first.

Tammy: No! Do 1 mL then 3 mL then 5 mL, and then 1 mL then 3 mL then 5 mL, and then 1 mL then 3 mL then 5 mL. Because if we do not have time to finish the last trial ... (Tammy's sentence is left unfinished however her idea is accepted.)

The following dialogue, during the Panadol investigation portrays how Kim and Tammy disagreed about the number of trials they should conduct. It seemed that Kim wanted to count the preliminary trials as the actual trials. Eventually, the issue was resolved when Gay sided with Tammy that they needed more trials.

Tammy announced, "We need to do six more boilings (trials)." Kim replied, "No! We only need to do four more (in total)." Tammy walked off annoyed, seemingly because they had reached an impasse. If they repeated the investigation twice then they would have a total of six trials to conduct. Kim didn't want to repeat

the two preliminary trials. She wanted to use them in her data collection. Gay hadn't arrived at the bench yet and Kim called her. Fifteen minutes later Tammy revealed that she had not surrendered her position on the number of trials. "Why don't we do two trials while we are waiting?" she said. To this Kim replied, "I've done them all. I've got them finished." Tammy then stated, "But we've not done them all!" At this stage Gay interrupted, "We need to do a boiling and room temperature." Tammy then said, "That's what I said. Then we would have done each trial three times. Yeah, that's for certain." At this stage Tammy departed to set up the additional trials.

In some ways this group functioned like groups in previously reported studies of laboratory work. For example, using the classification of roles described by Richmond and Striley (1996, p. 849) Tammy was the "leader", Kim most frequently the "helper" although she occasionally took a leadership role, and Gay vacillated between an "active contributor" and "passive non-contributor". The group focused on the procedures required to complete the task in a similar way to that reported by Christensen and McRobbie (1994), and they used "majority rule" to resolve difficulties (Roychoudhury & Roth, 1996, p. 431).

Scenario 9.

The entire investigation-related dialogue that occurred over 35 minutes is presented for Group TC. Their discussion focused predominantly on ways to complete the investigation and appeared unlikely to contribute to improved science conceptual understandings or investigation competencies.

Of the three groups, Group TC functioned the least effectively and least efficiently. They did not have an established leader and they had difficulty deciding what to do and how to do it. Olive appeared the most academically able, however, she tended to 'do her own thing' possibly because the other students had difficulties in completing tasks. For example, in the trypsin investigation she set up her own trials and stayed in the class on her own for the 20 minute recess period to complete the trialing. In contrast, Gemma and Jessica muddled through their trials not knowing the temperature at which Olive had collected her data. In the end, Olive collected data at two temperatures while Gemma and Jessica jointly performed one trial. Olive was respected by the others for her academic ability although at times she was aloof and did not relate well to them. This is supported by an interview

comment (p. 187) that she made about the difficulties in getting the group to agree with her. The inability of the group to organise themselves and to develop an overall plan is exemplified by the following dialogue. On two occasions when they did not know how to proceed they chatted socially. An example of parallel dialogue emerged when Olive decided to test her hypothesis and Gemma thought they were testing the hypothesis that she and Jessica discussed. Interestingly no one seems to take much notice of Gemma's ideas. She was overpowered by Jessica's scatty, enthusiastic but undirected approach.

Jessica: Olive, What's your hypothesis?

Olive: Body temp. works best.

Gemma: No! No! The change of temperature changes the time it takes to react. (Gemma and Jessica write this down.)

Olive: Guys what temperatures shall we do? (She ignores Gemma's comment.) I think we should do 30, 37 and 45 degrees. 26 is very humid (sic).

Gemma: We're meant to do a lower one.

Olive: Yeah, 30! (Olive has not listened to Gemma's hypothesis. She has selected temperatures almost evenly distributed around 37 °C)

Gemma: 30, 37 and 45. (Gemma writes this down.) I burnt my fingers in PVE (Personal and vocational education lessons).

They then discuss social issues for a considerable period of time. Much later they return to the task.

Olive: Guys, if we're sticking the film in water, then when does the trypsin acid come into it? Guys we have to work out when to put the trypsin acid in!

Jessica: What's the trypsin acid? What is that?

Gemma: That's the whole test sweetie pie. (Gemma rubs Jessica on the head.)

Jessica: I know, but do we put that in before the film or after the film.

Gemma: I don't get this.

Jessica reads out the instructions. Olive reads them to herself. Olive then disappears to ask another student and Gemma also disappears to ask the teacher.

Jessica: We don't need anything guys.

Jessica looks up from her reading but the others have gone and there is no one to tell what she has found out. She talks to another student about the security at her house. Two minutes later, after checking what students in other groups are doing, Olive arrives back at the bench with enthusiasm.

Olive: OK guys. We stick the trypsin in the test tube. We stick the film in …

Olive writes this down in her method section. Jessica doesn't listen. She is talking about a social event.

Jessica: What?

Jessica looks at Olive and sees that she is busy writing down something and then turns away to continue talking to her friend. Olive collects some equipment and four minutes later Gemma arrives back at the bench after checking with Miss Mills.

Gemma: Guys, we're doing this wrong!

Jessica: Are we? I thought this was right. OK. Let's read this guys.

Jessica starts reading the background information sheet from the beginning for the second time.

Gemma: What do we need the film for? (Miss Mills had asked them this in the introduction and they have been working on the task for 30 minutes, Scenario 6, p. 207.)

Jessica: 'Cos trypsenen (sic) reacts with, against, with, the stuff that's in the film.

Gemma doesn't listen. She's already moved off. They all move off and talk to other students in the class. Eventually Olive comes back with more equipment.

Jessica: So what are you doing now?

Olive: A preliminary trial to just see…

Gemma: Don't we need to time it?

Olive: No, not just yet. We stick it in the water bath.

They disappear to the water bath and four minutes later Jessica arrives back with Gemma.

Gemma: We don't need any water in the test tube do we?

Jessica No. We've got to change that to trypsin. (They have written 'water' on their worksheet.) She laughs, I mean enzyme.

Gemma: Yeah, 4 mL trypsin.

Jessica: No, it's enzyme!

Gemma: Well, that's the same then.

Jessica has not realised that the enzyme is trypsin. She thinks that they are different things.

Jessica: What did I say? Enzyme, yeah! I meant trypsin.

After the lesson has progressed for 35 minutes Olive arrives at the bench with the preliminary trial completed and announces, "That's what it's supposed to look like guys."

Jessica: We need to get some beakers to do 30 degrees and the other temps.

Olive: No! We get a beaker of 30 degrees and put **all** the test tubes in - otherwise it takes too long and we time them all together. Who wants to do... I'll do 45.

Jessica: I'll do 30 and Gemma you do 37.

They did not follow Jessica's instruction because Olive trialed at two temperatures and Jessica and Gemma only performed one trial together.

Members of this group appeared not to have specific roles as reported in the Richmond and Striley study (1996). They were the least organised of the groups studied. Similar to the Christensen and McRobbie study (1994), work-related conversation was predominantly focused on how to complete the investigation, although most of their conversation was about social issues and not documented here.

Scenario 10.

Group SC comprised Jody (the leader), Harue and Christie. Jody virtually harangued Harue and Christie to think about and complete the work. Harue made thoughtful contributions, although at times she seemed to be in her own world. Christie was less academically able than the others and she opted out of conversations that involved difficult concepts and engaged in social discourse with other students. She appeared most happy and comfortable when doing the Panadol investigation, possibly because this was the most straight forward investigation. The following conversation reveals Jody's organising ability and the willingness of the others to comply with her leadership. It also reveals Christie's tendency to opt out of difficult concepts.

Jody: Guys, let's all concentrate this time.

Christie: I'm trying to work out something.

Jody: What?

Christie replies "Nothing," and she adds in a pathetic voice almost as if abrogating herself of any further responsibility, "I don't understand." She then turns to talk to a student from another group.

Harue: What variables shall we look at? Its acidity and the amount of temperature are the only ones I can think of. We can vary the temperature. The basic solution (in which the trypsin reacts), we could make it acidic.

Jody: We don't have the materials to make it acidic or basic.

Harue: She (the teacher) said we had hydrochloric acid.

Jody: Oh, did she! The amount of trypsin, but I don't think we should do that.

Harue: Well, do you want to do the temperature?

Jody: Yeah, we'll do that. Guys, concentrate please!

Christie is still talking to the other student and has not participated in the decision made by Jody and Harue. She says, "I'm just trying to work something out." To which Jody replies, "What?" Christie then repeats for the second time, "I don't understand." Jody and Harue continue to discuss the temperature.

Jody: Well, it (the change in temperature from the optimal conditions) would be just as bad either way because it's got to be close to 37 degrees doesn't it? 36 and 38, it will be equally bad either way.

Harue: But we've got to say it will be better if it's hotter or colder.

Jody: We'll need to keep it close (to 37 °C) to show that it's that one (the optimal temperature).

They ask Mrs Grant and she says that she wants them to work it out themselves. Jody starts to write up the method. "Hey, hey!" says Harue. Let's get the hypothesis first." Jody replies, "Well, we'll work down through it (the IPRS)." Jody starts to dictate, "The further the temperature gets from 37 the longer it takes. But we have to improve on the wording." They read from the instructions that the optimum conditions are with 4 mL of trypsin and 1 mL of sodium hydroxide. They think they need some water, however, they are not sure what to do with it. They list what they need and decide to gather data at temperatures of 36, 37 and 38 °C. They decide to do nine trials, three for each temperature. Harue persists, "Why do we need any water?" Jody replies, "Well, to put it in a bath." They continue writing an equipment list and Jody says, "OK, let's go through it and see if we've got what we need." They add safety glasses and gloves. As the bell goes Jody says "OK guys, just think about it so that when we come back we've got ideas and everything."

The next lesson commences and Jody says, "I think we should change to the amount of trypsin (instead of investigating the effect of temperature on the reaction). It would be a lot easier." Christie and Harue both respond "OK." They check with the group on the next bench and they are also doing the amount of trypsin. "Yeah,"

says Jody, "otherwise we will be fiddling around with the amount of ice all the time trying to get the right temp. This will be a lot easier." Christie grumbles, "Now I have to change everything." To which Jody responds, "That's why I do mine in pencil." They then decide to time until half of the film becomes clear. Jody sends them off to collect the equipment. She stands at the bench and checks that they return with everything. Then, while the others are not present she says to the camera, "I love bossing people around." Harue comes back and says, "OK guys, let's get started." Jody says, "No, do the table first." But this time Harue wins out and they do the preliminary trial. They trial with 2 mL of trypsin and discover that they don't really know what they are measuring. They ask Mrs Grant and she says, "What is the reaction?" and "How can you tell when it's reacted?" Harue answers both questions correctly and Jody is peeved that she couldn't answer. "I hate it when I don't know what I'm supposed to be doing, and no one else knows, and they all start telling me different things, and I keep doing it wrong, and then everyone blames me for it," she says. "We're not blaming you," says Harue. The bell goes and they finish off the investigation during the next lesson.

The roles of group members align with Richmond and Striley's (1996, p. 849) classification with Jody the "leader", Harue the "helper" and Christie a "passive non-contributor". Although the group's decision to change from investigating temperature to the amount of trypsin was based on the relative ease of the procedures, Harue and Jody engaged in productive dialogue about the variables to investigate and optimal conditions. In contrast, Christie made very little contribution to the group.

In summary, each group engaged in a considerable amount of social discussion. At times these discussions appeared to be triggered when there was some doubt about how to proceed with the investigation. This tendency was particularly evident with Scenario 9 and with Christie in Scenario 10. These ideas are expressed as Assertion 68.

Assertion 68: During investigations students engage in discussions about social issues, especially when they encounter difficulties with the investigations.

Group discussions about laboratory investigations predominantly focused on ways to accomplish the task and this notion is expressed as Assertion 69. From the audio and video data, there was little evidence of dialogue that might contribute to the development of improved conceptual understanding and investigation competencies (Assertion 70). This finding appears to be in conflict with an earlier assertion (Assertion 25) that "Most students perceive that they learn more about doing investigations from talking with their peers than from talking with their teacher." Based on the audio and video data, however, an appropriate interpretation would be that students learned how to complete the investigation from their peers but they were unlikely to improve their conceptual understandings or investigation competencies from the interactions with their peers. This has parallels with the work of Christensen and McRobbie (1994) who describe laboratory work in which students' main focus was on the physical operations of the task, without discussion of the purposes of the task or its conceptual meaning. Surprisingly they found that students "believed quite strongly that practical work was vital to their understanding of concepts" (p. 58). In the Christensen and McRobbie study, and also with this research, the mechanisms by which practical work contributes to learning are not clear. Students in this research perceived that learning occurred and gains in investigation competencies measured by pretest and posttest data confirm this perception. Kempa and Ayob (1995) contend that task related comments and observations made by students during group discussion represent a major shared knowledge resource from which students learn. They also contend that a significant portion of information appearing in students' written responses to questions was neither prior knowledge nor part of the shared knowledge. They make no claims about how students' information/understanding was constructed. Hence, while the

observational data suggest superficial and low level learning, it may be that observations per se are too superficial to access and measure relevant information about how understandings are generated in group settings.

Assertion 69: Group discussions about investigations are mostly concerned with how to complete the investigation.

Assertion 70: Few group discussions appear to have the potential to improve students' understandings of concepts or to develop investigation competencies.

Assessment

Three scenarios are presented which describe the implementation of teacher assessed norm referenced, teacher assessed criterion referenced and student assessed criterion referenced assessment in Classes TN, TC and SC respectively. Similarities and differences in the implementation of these regimes are discussed and a number of assertions are made.

Teacher Assessed Norm Referenced Assessment

Scenario 11.

Following the assessment of their investigations, Class TN received verbal feedback from their teachers for approximately 20 minutes. With the Panadol investigation, Mrs Cross went through the IPRS and advised students of the correct answers and highlighted common errors on the white board. She was business like in her approach and did not engage in drawn out explanations. She mentioned that the graph was poorly done because the scale was not accurately drawn, and said that it was not appropriate to say the temperature was hot or cold when a thermometer

was provided. She also mentioned that preliminary trials should be used to see how much water to use; to practise the procedures; to find out what temperatures worked best; to check the timing; and to define what 'dissolve' meant. Mrs Cross said that the sample size of three was tiny and because of this students should not be very confident in their findings.

After each investigation the students paid little attention to the teacher's verbal feedback. This is exemplified by the actions of Kim, Gay and Tammy with the Panadol investigation feedback. Gay spent the entire time drawing on her file and Kim started to fill in the IPRS for the following investigation. Tammy spent some time making corrections, and then together she and Kim wrote up the next investigation.

In terms of the written feedback on the students' IPRS, Mrs Cross assigned each member of the group an overall B grade for the Panadol investigation. She wrote specific comments throughout the students' reports such as, "Graph needs a title," and more general comments at the end of the reports, for example, "You haven't completed a lot of the sections in enough detail. For example, there is no diagram or plan. Evaluating the findings was poorly done. It is important to include all these sections in your investigation." It is interesting to note that for this investigation the group plotted temperature as a discrete variable (ice, room temperature, water bath and boiling) albeit, during the lesson, the teacher had told them to be more accurate and to record the actual temperatures (Scenario 5, p. 208).

The students' acknowledgment of the teachers' (Mrs Cross and Mrs Grant) written and verbal comments was minimal. For the target group the only evidence of students making corrections in response to feedback was Tammy's IPRS for the Panadol investigation. When Mrs Cross provided verbal feedback in class, Tammy added some information to the planning section of her report, such as how to change and measure the variables and the need to include a diagram. All of the IPRSs for the class were examined and only a few additions or alterations were made as a result of the feedback.

Teacher Assessed Criterion Referenced Assessment

Scenario 12.

Teachers of Class TC devoted approximately 20 minutes to verbal feedback in the whole-class setting and then allowed students 10 minutes to make corrections. In the feedback provided by Mr Brogo (electromagnet investigation) he said that the students had improved in their descriptions of the investigation and mentioned other areas where they could improve, including table headings and relating the conclusion to the hypothesis. He added that they were overly confident in their findings. He talked about errors arising from the measurement of the dependent variable (the attraction of the paper clip to the electromagnet), the fact that the resistance of the solenoid increased as it became hot, and that they should attempt to explain anomalous results. He added that, "with real science you are working in the dark because you don't know what the result will be," and explained that in earlier times scientists made the equipment before they could collect the data. During the lesson many students were inattentive. For example, Mr Brogo asked students to raise their hands if they had a constant reading for the dependent variable after changing the independent variable. Only one out of a possible six students raised her hand. He then complained to the students that they were not listening. The audio and video recording showed several students quietly talking and some were touching each other with a ruler.

When Group TC received their assessed IPRSs they first tallied the number of A, B and C grades that they had been allocated. After an initial flurry of interest they rapidly became off-task. Their behaviour was within the bounds of accepted and appropriate school classroom behaviour so they were able to escape the teacher's attention, however, the audio and video recording captured their perfunctory responses. Jessica, for example, spent the whole session discussing her teeth. She showed everyone the "separators" in her mouth and said that she was soon getting braces.

Evidence for Group TC's minimal response to the feedback is also captured throughout their worksheets and IPRSs. For three out of four investigations, teachers had written that they had not adequately stated how to measure the dependent variable. This omission was repeated in their posttest investigation, clearly indicating that they had not acted on the feedback. Also, for each table of results the students drew, their teachers commented that they had omitted information (a title and reference to the dependent variable) and for their posttest investigation they also omitted this information. As well, they treated seconds as

decimal places of a minute to obtain an average of 15 minutes and 67 seconds, which provides further support for Assertion 9.

Another consistent and unresolved difficulty for Group TC was dealing with continuous data. For the ukulele investigation they should have plotted the frequency of the note against the length of the string, however, they rendered their independent variable discrete by plotting the fret number. Hence, the longest string was fret 1 and the shortest was fret 11. Their graph showed odd numbered frets, Fret 1, Fret 3, Fret 5 and so on, evenly spaced on the horizontal axis. For their graph Mr Brogo wrote, "It would have been better to plot string length." For the electromagnet investigation, their horizontal axis was correct and this may have been due to the good fortune that they had increased the voltage uniformly by 2 volts each time from an initial reading of 2 volts to 12 volts. Therefore, the axis went up evenly. Scenario 7 described their posttest graph in which they made errors determining the axis for continuous data. Problems that student experience in viewing continuous data as discrete data were also evident from the IPRS data and were expressed as Assertion 17.

Student Assessed Criterion Referenced Assessment

Scenario 13.

Class SC was allocated 30 minutes to conduct the assessment and make corrections to their IPRSs. Group SC applied a consistent procedure in which Jody took control. They turned to the beginning of their IPRS and Jody read out the answers from the master answer sheet. She said, "Yeah, we got that," for most of their responses. Christie and Harue didn't look at the master sheet. Harue followed Jody's instructions. For some investigations Christie did not assess her work. Mostly Jody and Harue ticked their responses; they occasionally added extra information and they seldom crossed out their responses. There was very little discussion. They then turned to the page on which the criteria were listed and assigned grades for each criterion. For most of their work they graded themselves A and when they had experienced problems doing some aspect of the investigation they assigned a B grade. Often, however, they had solved their problems by the end of the investigation and could have been more generous in determining a summative grade. For those criteria where they had insufficient data they assigned themselves a C or NR grade. They finished the task and spent about 15 minutes discussing sport and the school camp arrangements. They talked quietly and did not attract the teacher's attention.

Similarities and Differences between the Assessment Regimes

The intended implementation of the assessment regimes is described previously in Chapter 3. Instructions given to teachers about the assessment procedures are presented in Appendix I. The actual implementation was consistent with the intended implementation and is summarised as follows.

- For **Class TN**, teachers wrote comments throughout the IPRSs and **a final comment** on the last page where the overall grade was assigned. To determine the grades teachers ranked the IPRSs and allocated grades as described in Appendix I. Teachers spent approximately 20 minutes going over students' weaknesses in a whole-class setting. Following this students were allocated 10 minutes to make corrections based on the teacher's feedback.
- For **Class TC** teachers wrote comments throughout the IPRS and **brief comments for the criteria**. They assigned grades for each criterion. The teachers spent approximately 20 minutes going over students' weaknesses on the IPRS in a whole-class setting. Following this students were allocated 10 minutes to make corrections based on the teacher's feedback.
- In **Class SC** the students were allocated 30 minutes to check their IPRS with a master answer sheet. They assigned a grade for each criterion and made corrections to their work.

The implementation of the two teacher assessment regimes was very similar (Assertion 71) and the student assessment regime was different. For Class TN and Class TC there was no observable difference in the quantity and quality of the verbal feedback. The only difference between Classes TN and TC was the form of the written feedback: Class TN had an overall comment and Class TC had comments that addressed the criteria. In effect, there was no real difference because the teachers of Class TN used the IPRS to provide the structure for their comments. For example, Mrs Grant listed the question numbers for the areas where improvements

could be made, and Mrs Cross wrote general comments about planning investigations, conducting investigation, processing data and evaluating the findings. Hence, they addressed the same information that was addressed.

Assertion 71: The implementations of teacher assessed norm referenced assessment and teacher assessed criterion referenced assessment were very similar.

For Class TN and Class TC teachers provided written feedback throughout the IPRS by identifying areas of weakness. They did this by

(a) crossing out incorrect words and writing the correct word over the top. For example they wrote the 'number' of pins instead of 'amount' of pins.

(b) requesting explanations based on scientific language or understandings. For example when students were asked to write **why** they could draw a conclusion their response was, "because each time we uped the vaults (sic) more paper clips were picked up, proving that the magnetic field had gotten wider, could cover a greater distance." To this the teacher responded "Does not explain why **volts** gives us a stronger magnetic field." Another example occurred when the student wrote "so that the data could be deadly accurate," and the teacher responded that she should not used terms like **deadly** because they were not scientific.

(c) identifying omitted information. For example they wrote, "What are these values?" or a question mark for tables that lacked a heading.

(d) requesting students to expand on their responses by asking for more information. For example teachers wrote, "Far more detail is needed. Think a little harder regarding the possible limits of the experiment." Also when a students wrote that an error in their investigation arose because the "power magnet did not work properly," the teacher replied by saying that she was not

sure what was meant by the statement and asked the student to explain in more detail.

(e) identifying irrelevant information by crossing it out and stating that it was not needed.

(f) praising or ticking correct information.

The Group TN and Group TC, (Scenarios 11 and 12), paid little attention to the teacher feedback on their investigation performance. This notion has been expressed as Assertion 72. Further an examination of their IPRSs showed that most students did not make any written corrections to their work (Assertion 72) during the ten minutes allocated for them to make corrections.

Assertion 72: When receiving verbal teacher feedback about investigation performance, most students are inattentive in whole-class settings.

Assertion 73: Following teacher feedback about investigation performance, most students do not make written corrections to their Investigation Planning and Report Sheets.

The implementation of the student assessed criterion referenced assessment differed from the teacher assessments. Although the students' involvement in the task was necessarily greater than for the assessment procedures in the other classes, it was superficial. When Jody, Harue and Christie assessed their work (Scenario 13) it seemed as though they focused on confirming their answers rather than identifying ways to improve their performance. It was evident, also, that they lacked opportunities for high quality feedback that could be obtained from their teachers and this supports Assertion 58. A number of difficulties that students experienced with this assessment are discussed.

One problem that arose was due to the mismatch between the students' science vocabulary and that of the teacher. Jody, the leader said, "Let's go through it (the master answer sheet on the electromagnet investigation) and look at it all and then do our A, B and Cs and things." They matched up their answers with those on the master sheet and ticked or crossed as they proceeded. In their list of variables they crossed out their responses, **the size of the iron bar** and **the amount of iron bar** because the master answer sheet said **the amount of core in the solenoid**. "Gee! We got totally different stuff," said Jody. "Well, we did (conducted) it right!" she added emphatically. Clearly, if the teacher had assessed their investigation then they would have accepted the students' responses. This is expressed in Assertion 74.

Assertion 74: Some students using student assessed criterion referenced assessment, assume their responses to questions are incorrect because their language differs from that on master answer sheets.

A second problem arose because students were unable to access the teachers' responses because of inadequate science content knowledge (Assertion 75). For example, in the list of variables the master answer sheet mentioned the variable "temperature" because at low temperatures the current would be greater and hence the magnetic effect would be greater. "I don't think that has anything to do with it," said Christie. "It's got (on the master answer sheet) use a switch so the circuit doesn't get too hot," she added in a puzzled tone. They ignored this information because they did not have the science background knowledge to make meaning from the information. (In contrast Mr Brogo addressed this point in the feedback that he gave to Class TC, Scenario 12.)

Assertion 75: Some students using student assessed criterion referenced assessment, are unable to access information on master answer sheets because they lack adequate science background knowledge.

A third problem was the failure of students to recognise all the relevant feedback cues (Assertion 76). As a result, they did not consistently identify all the areas where improvements could be made even though the information was presented to them on the master answer sheets. For example, for each results table they drew, Group SC did not include a heading for the measurement of the dependent variable even though this was present on the master answer sheet. In contrast some feedback was acknowledged. For example, when they read the master answer sheet for the electromagnet investigation, they realised that they had not considered the friction between the paper staple and the slightly uneven bench top as a source of error.

Assertion 76: Some students using student assessed criterion referenced assessment fail to identify all the relevant feedback on master answer sheets.

Summary of the Chapter

This Chapter has discussed the audio and video data collected throughout the instructional program. Twelve assertions have been presented about the cognitive apprenticeship model of instruction and six have been presented about the different assessment regimes (Figure 21). Of the assertions about the cognitive apprenticeship model of instruction, five were associated with teacher modelling,

two with scaffolding and fading, two with teacher guidance, and three with articulating.

It was found that the implementation of the teacher assessed norm referenced assessment and the teacher assessed criterion referenced assessment were very similar. Therefore three assertions refer to both of these assessment regimes. Three assertions are associated with the student assessed criterion referenced assessment.

In this Chapter, previously established Assertions (9, 17, 40, 41, 42, 44, 53 and 58) were supported by the audio and video data. Assertion 25 which was generated from student questionnaire data: Most students perceive that they learned more about doing investigations from talking with their peers than from talking with the teacher, contrasts with Assertions 68, 69 and 70 generated from the audio and video data. These assertions (68, 69 and 70) suggest that student discussions are largely social, that they focus on how to complete the investigation, and that there is little discourse that might develop science conceptual understandings or investigation competencies. These assertions need also to be considered in light of pre and posttest data that indicates significant gains in students' investigative competencies (Chapter 4). One interpretation is that observational data is a poor indicator of students' learning. Another is that students interpreted the phrase 'doing investigations' on the questionnaires at a very superficial level and associated it with 'completing the task'. If this were the case then Assertion 69: Most group discussions that address science investigations, focus on the procedures required to complete the task, may be seen as supporting Assertion 25.

Assertions
Aspects of the cognitive apprenticeship model of instruction
(a) Teacher modelled investigations
59 During teacher modelled investigations on task student behaviour is maintained by requiring students to complete Investigation Planning and Report Sheets
60 Different investigations provide opportunities for teachers to model different investigation competencies
61 Second hand data can be used effectively to model data processing competencies
62 Moving students from their seats so that the data gathering can be observed may result in increased off-task behaviour
63 Students appear to lose interest when watching repeated or replicated trials
(b) Scaffolding and fading
64 Teachers use the scaffolding provided by Investigation Planning and Report Sheets to organise instruction and to provide a focus for student learning.
65 Students use the scaffolding provided by Investigation Planning and Report Sheets to ensure that they complete the requirements for investigations.
(c) Teacher guidance
66 Teachers differ in the type and amount of guidance they give students performing investigations
67 Teachers do not always gather appropriate information about students' understandings to make informed decisions about ways to coach students
(d) Articulating
68 During investigations students engage in discussions about social issues, especially when they encounter difficulties with the investigations
69 Group discussions about investigations are mostly concerned with how to complete the investigation.
70 Few group discussions appear to have the potential to improve students' understandings of concepts or to develop investigation competencies.
Assessment
(a) Teacher assessment
71 The implementations of teacher assessed norm referenced assessment and teacher assessed criterion referenced assessment were very similar.
72 When receiving verbal teacher feedback about investigation performance, most students are inattentive in whole-class settings.
73 Following teacher feedback about investigation performance, most students do not make written corrections to their Investigation Planning and Report Sheets.
(b) Student assessed criterion referenced assessment
74 Some students using student assessed criterion referenced assessment, assume their responses to questions are incorrect because their language differs from that on master answer sheets.
75 Some students using student assessed criterion referenced assessment, are unable to access information on master answer sheets because they lack adequate science background knowledge.
76 Some students using student assessed criterion referenced assessment fail to identify all the relevant feedback on master answer sheets.

Figure 21. Summary of the assertions developed from audio and video data

CHAPTER 9

DISCUSSION AND GENERALISATIONS

Overview of the Chapter

The Chapter draws together data presented previously in Chapter 4, (the Test of Science Investigation Skills, TOSIS, and Investigation Planning and Report Sheets, IPRS), Chapter 5 (student questionnaires), Chapter 6 (teacher interviews), Chapter 7 (student interviews), and Chapter 8 (audio and video recordings). Consistent with the earlier chapters, assertions generated from the data are clustered into the three themes that underpin the research; investigation competencies, the cognitive apprenticeship model of instruction, and assessment.

Schematically the clustering of assertions is presented in Figure 22, and this represents an extension of Figure 2. The assertions are coded according to the data source from where they originated with the Investigation Planning and Report sheet coded 'I'; student questionnaires coded 'Q'; student interviews coded 'S'; teacher interviews coded 'T'; and the researcher's interpretations of the audio and video data coded 'V'. Assertions that are triangulated from different data sources have codes, representing the various data sources, with the first mentioned code referring to the data source from where the assertion was generated. The process of triangulating data is considered in Chapter 3 to be an alternative to data validation (Denzin, 1989a, 1989b; Fielding & Fielding, 1986; Flick, 1992). This does not imply that data that are not triangulated are not valid, but implies that this method of analytical induction does not handle rare events well (Erickson, 1986).

From the clusters of assertions "key linkages" and "patterns of generalisation within the case at hand" are used to formulate general assertions (Erickson, 1986, p. 148). The general assertions are designated with the letters of the alphabet (Figure 22) and the assertions on which they are grounded are numbered. In addition, some general assertions were based on quantitative data gathered from pre and posttests (General Assertions M, N, and O) and one was based on the researcher's interpretations of the conventions and procedures of scientific inquiry (General Assertion E).

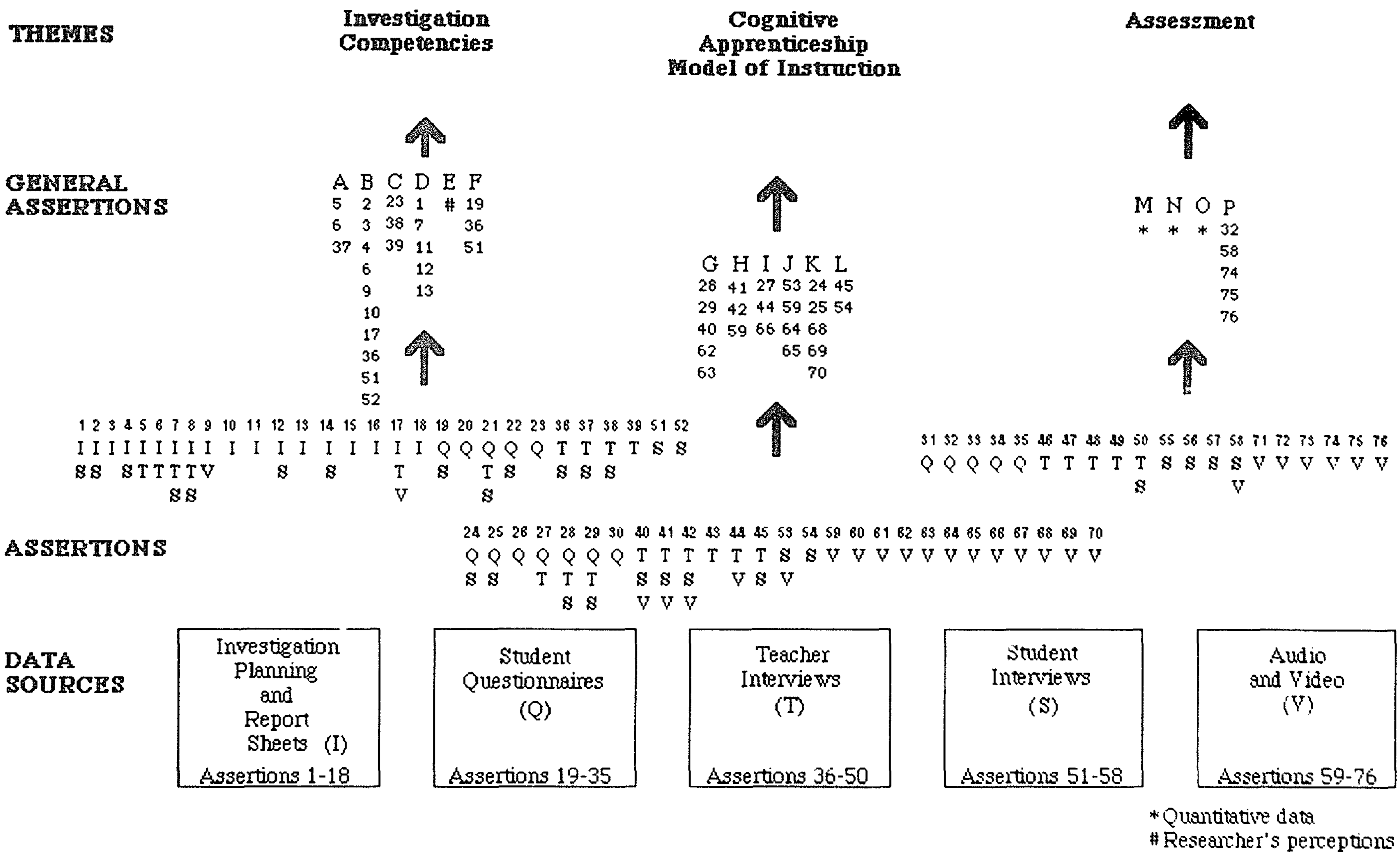


Figure 22. Schematic representation of the association between the data sources, assertions, general assertions and themes of the research

Investigation Competencies

Previously presented pretest and posttest data for the TOSIS and IPRS (Chapter 4) revealed statistically significant improvements in students' investigation skills and competencies resulting from the instructional program. From the pencil and paper TOSIS, a two-way Analysis of Variance (ANOVA) of students' total scores indicated an effect for test occasion of $F(1, 2) = 66.68$, $p < .01$. Descriptive statistics for the total score are presented in Table 1 and in Figure 6. Descriptive statistics for specific skills or concept areas assessed by the TOSIS demonstrated students' improvements in *Identifying variables, Writing an hypothesis, Planning an investigation, Drawing conclusions, Identifying methodological limitations,* and *Understanding the concepts of hypothesis, theory, data and conclusions.* These are presented in Table 2 and Figures 7 to 12.

The IPRS assessed students' performance of a science investigation in a more holistic and authentic manner than the TOSIS. The Wilcoxon Matched-Pairs Signed-Ranks test indicated that for each of the three groups/classes, the posttest levels of performance were significantly greater than the pretest levels ($p < .01$) for *Planning investigations, Conducting investigations, Processing data* and *Evaluating investigations.*

Through questionnaires and interviews, students reported on the competencies they perceived they had learned and these are expressed as Assertion 22: Investigation competencies that students claim to have learned include (a) planning investigations that comprises writing hypotheses, identifying variables, overall planing and preliminary trialing that can result in modifications to the plan (b) conducting investigations that comprises using equipment, controlling variables, repeating trials, being accurate and working safely (c) processing data that

comprises using tables, averaging results, using graphs and drawing conclusions (d) writing reports (e) working cooperatively (f) attending to detail (g) managing time and (h) being organised. In the main they reported that they had learned isolated process skills rather than associated or linked skills such as developing conclusions based closely on the data and related to the hypothesis. In addition, they reported learning several competencies which are usually considered to be social and workplace skills which are acquired through group learning (Linn & Burbules, 1993) such as *Working cooperatively, Managing time* and *Being organised*. As well, they recognised the quantitative nature of investigations and said that *Being accurate* and *Attending to detail* were learned.

The assertions associated with investigation competencies that were generated from the IPRS, student questionnaires, teacher and student group interviews and audio and video data are loosely grouped under the headings, planning investigations, conducting investigations, processing data, evaluating investigations, and other assertions in Figure 23. Many of the assertions are concerned with difficulties that students experienced when performing investigations. Six general assertions (A to F) are proposed from these data (Figure 22) and are addressed in the ensuing discussion.

Assertions

Planning investigations

1 Many hypotheses that students investigate are unlikely to be supported by data [I, S]

2 Many student think that an increase in temperature decreases the time taken for all reactions (including reactions involving biological enzymes) [I, S]

3 Many students interpret the time taken for the reaction to mean the time for the reaction to go to completion. This results in difficulties when they collect their data over a fixed time period and hence measure the rate of the reaction [I]

4 Some students who investigate the effect of the amount of substance in a chemical reaction on the time taken for the reaction, do not control for the volume of solution [I, S]

5 When planning investigations most students use preliminary trials to observe the reaction experiment and to organise procedures [I, T]

6 When planning investigations most students do not consider the range over which the data should be collected or the data collection interval. [I, T]

37 Most students' written plans for investigations lack sufficient detail to describe the procedures they intend to follow. [T, S]

Conducting investigations

7 Some students 'observe' changes that they think will occur in investigations rather than changes that actually occur. [I, T, S]

8 Most students learn to repeat or replicate the data collection and to average the results. [I, T, S]

9 Some students treat time as a decimal measure and consider seconds as one-hundredths of a minute. [I, V]

10 Most students do not have an appreciation of the degree of accuracy to which measurements should be recorded. [I]

38 Students' familiarity with the dependent variable, and how it is operationalised and measured affects the difficulty of investigations. [T, S]

39 Students find investigations that involve setting up equipment more difficult that those that do not require equipment to be set-up. [T]

Processing data

11 Some students confirm their hypothesis with insufficient data. [I]

12 Some students confirm their hypothesis with data from dubious sources. [I, S]

13 Some students ignore or reject data so that they can confirm their hypothesis. [I]

14 When faced with unexpected results some students reject their original hypothesis and formulate a new conclusion based on inconclusive evidence. [I, S]

15 Some students collect data in order to confirm a different hypothesis from that which they originally planned to investigate [I]

16 Some students investigate an hypothesis that cannot be supported and use their data to reject the hypothesis [I]

17 Some students view continuous data as discrete data [I, T, V]

Evaluating investigations

18 Most students perceive that it is necessary to investigate and confirm a scientifically 'correct' hypothesis [I]

21 Students do not consider *Evaluating investigations* as learning about investigation competencies [Q, T, S]

Other assertions

19 Students perceive that they learn most from the first few investigations that they perform. [Q, S]

20 Students perceive that they learn most about writing hypotheses and identifying variables from the worksheet on terminology. [Q]

22 Investigation competencies that students claim to have learned include (a) planning investigations that comprises writing hypotheses, identifying variables, overall planning and preliminary trialing that can result in modifications to the plan (b) conducting investigations that comprises using equipment, controlling variables, repeating trials, being accurate and working safely (c) processing data that comprises using tables, averaging results, using graphs and drawing conclusions (d) writing reports (e) working cooperatively (f) attending to detail (g) managing time and (h) being organised. [Q, S]

23 The nature of the investigation influences the investigation competencies that students believe they learn. [Q]

36 Teachers consider that some students do not perceive developing investigation competencies to be learning science. [T, S]

51 Most students prefer to investigate science problems for which they do not know the answer/s. [S]

52 The difficulty of investigations is influenced by students' prior conceptual knowledge associated with the investigations. [S]

Figure 23. Assertions about the development of investigation competencies

Note: I = Investigation Planning and Report Sheet, S = student interviews, Q = student questionnaires, T = teacher interviews, and V = audio and video data
The code first mentioned indicates the source of the assertion. Subsequent codes indicate triangulation of the data from other sources.

Difficulties experienced by some students appeared to arise from their inabilities to form a **conceptual model of the investigation** in the planning phase. These students lacked a sense of direction in the investigation and they were unable to identify and focus on important relevant details. For some students the consequence was that they did not complete the investigation. The lack of direction was particularly evident from the pretest where 24 out of 65 students (37 %) did not complete the *Evaluating investigations* section of the IPRS (Table 3). Lack of attention to relevant details is evident from the following assertions; Assertion 37: Most students' written plans for investigations lack sufficient detail to describe the procedures they intended to follow; and Assertion 6: When planning investigations most students do not consider the range over which the data should be collected or the data collection interval.

The importance of forming a conceptual model of the process to be learned is documented by Lave in Collins, Brown and Newman (1989). Lave contends that a conceptual model serves three purposes. First, it provides an advanced organiser to allow learners to concentrate more of their attention on the execution of the process than would otherwise be possible. Second, it provides an interpretative structure for making sense of the feedback, hints and corrections during interactive coaching sessions, and third it provides an internalised guide for independent practice. Most relevant to this discussion is the purpose of enabling the learner to concentrate more on the execution of the task. Hence, implicit in this notion is that when students have formed a conceptual model of an investigation they are more likely to attend to relevant details. A secure conceptual model may serve to reduce information overload and behaviours such as concentrating on one part of the experiment and excluding the rest, and busy random activity (Johnstone & Wham, 1982) may not be as prevalent when students have a conceptual model of an investigation. This notion is supported by research (Chi, Feltovich, & Glaser, 1981; Hackling & Garnett, 1995) that highlights domain specific schema knowledge as an influence in successful problem solving.

Without this schema, Year 7 students performing science investigations lacked direction and did not focus on details (Hackling & Garnett, 1995), similar to students in this study. In comparison, expert research scientists work from a schema, exhibit high levels of metaplanning (Hackling & Garnett, 1995; Hayes-Roth & Hayes-Roth, 1979), and are able to focus on details. These ideas are presented as follows.

General Assertion A: When students do not have an appropriate conceptual model of science investigations they lack direction, they do not pay attention to relevant details and/or they often do not complete the investigation.

Several assertions are associated with **science conceptual knowledge**. For example, Assertion 2 (the relationship between temperature and time for an enzyme-catalysed reaction) shaped students' expectations of the experimental results. Many students expected the enzyme activity to increase with temperature and obtained confirming data even though this is contrary to established science. Assertion 3 (the difference in meaning between the rate of reaction and time for reaction to go to completion) is a complex issue and possibly beyond the understanding of many Year 9 students. The difference betwe n these concepts led to difficulties in identifying an appropriate measure of the dependent variable. These assertions indicate that limited science conceptual knowledge can affect success at performing investigations. The science conceptual knowledge that underpins the task, and students' understandings of this conceptual knowledge affects students' success at investigating (Coles & Gott, 1993; Germann & Aram, 1996; Gott & Duggan, 1995). Care must be exercised, therefore, to ensure that the difficulty of the investigation is at an appropriate level.

During interviews, teachers said that some students did not perceive performing investigations to be learning science (Assertion 36) and implied that learning about the way science is conducted is perceived by students to be of less

value than learning about science concepts. This is supported by the fact that most students preferred to investigate problems for which they did not know the answers (Assertion 51) presumably so that they could develop conceptual understandings. They commented that the Panadol and ukulele investigations were "general knowledge" because they were familiar with the relationships being investigated. These investigations were perceived to be easier than the others and Assertion 52 is based on the premise that the difficulty of an investigation is influenced by a student's prior conceptual knowledge about the investigation. These assertions affirm the need to link investigations with science conceptual knowledge, a point that is made by Coles and Gott (1993).

Students' **knowledge and application of procedural skills** also influences investigation performance (Coles & Gott, 1993; Germann & Aram, 1996). In this research the following assertions about procedural knowledge were generated; Assertion 4 (controlling variables), Assertion 6 (data range and interval), Assertion 9 (recording time), Assertion 10 (accuracy of measurement), and Assertion 17 (continuous and discrete data). Difficulties associated with graphing discrete and continuous data have been described previously by researchers, Beichner (1990), Roth and McGinn (1997), and Wavering (1989).

Another shortcoming with some students' performances was that they did not quantify their data in circumstances where quantification was appropriate. Examples were evident when students indicated that their test tube was "one quarter full"; when Tammy said the temperature was "boiling"; when Group TC indicated the length of the ukulele string by the fret number; and when Jo, Janet and Ed indicated that their temperatures were icy, room temperature and boiling. Black (1990, p. 21) believes that students only quantify information if they perceive that quantification is a "powerful tool" in science. Reif and Larkin (1991) state that accurate quantification is a difference between science and everyday life because in

day to day life students do not have to be extremely accurate and it is likely that previously students may not have needed to make accurate measurements.

The preceding discussion and the clustering of several assertions leads to the generation of a more general research finding. The fact that students' science conceptual knowledge, and their knowledge and application of skills and procedures influence their abilities to perform investigations is expressed as General Assertion B.

General Assertion B: Students' abilities to perform investigations are influenced by their understanding of the science content knowledge that underpins the investigation, and their knowledge and understanding of investigation skills and procedures.

In addition to students' prior understanding of science content knowledge and skills and procedures, factors associated with **the nature of the task** also influence their abilities to perform investigations and the notion that different investigations develop different investigation competencies is expressed as Assertion 23. Difficulties students experience in this study are associated with operationalising and measuring the dependent variable (Assertion 38) and with the complexity of setting up the equipment (Assertion 39). Sometimes the particular variable that students chose to investigate affected the difficulty of the investigation. For example, temperature was more difficult to control and monitor than the amount of catalyst (Scenario 10). Also, the independent variable which was selected determined whether a line or a bar graph was required, and students found line graphs more difficult than bar graphs because they need to determine a scale for the independent variable (Scenario 7). For example, when temperature is the independent variable a line graph is needed to display the results, however, when the nature of the solvent is selected a bar graph is required. Findings about how the

nature of the task affects the difficulty of an investigation are presented in General Assertion C.

The findings of this research about the nature of the investigation affecting its difficulty are supported by other research. The following factors have been identified as influences in students' success at performing investigations; whether the context is an everyday context or a science context (Gott & Duggan, 1995); the clarity of the task (Germann & Aram, 1996); the complexity of the task including the number and type of variables (Coles & Gott 1993; Gott & Duggan, 1995); the potential for the task to result in information overload (Johnstone, 1980; Johnstone & Letton, 1990, 1991; Johnstone & Wham, 1982) and the interaction of the science concepts and the procedural complexity (Gott & Duggan, 1995). In addition, student attributes have been the focus of other research which reports the following influences on students' abilities to perform investigations; the age of the students (Gott & Duggan, 1995), their commitment and determination (Gott & Duggan, 1995; Hodson, 1992), and their abilities to associate different aspects of the investigation, such as the conclusion to the hypothesis (Hodson, 1992).

General Assertion C: The difficulty of an investigation is influenced by factors associated with the nature of the task such as the complexity of setting up the equipment, the type of variables (whether they are discrete or continuous) and way the variables are measured and controlled.

Often students were unable to view data objectively and several assertions were concerned with **the ways students confirmed existing beliefs**. These assertions appeared to emerge as a consequence of students proposing hypotheses that were not supported by data (Assertion 1). The result was that students attended to their data in such ways as to confirm their hypotheses (Assertions 7, 11, 12 and

13). The tendency of some students to seek confirmatory evidence and their failure to acknowledge disconfirming data may indicate that they are certain that their beliefs (original hypotheses) are correct and that these beliefs are resistant to change. This finding is expressed in General Assertion D.

This finding has parallels with research conducted by Kuhn (1992), Novak (1988), and Nussbaum and Novick (1982). Kuhn noted that people across a range of ages tend to hold theories with certainty and that from one half to three quarters of a sample of 160 people claimed that they were sure or very sure that their theories were correct. From studies of students' conceptions, Novak, and Nussbaum and Novick contend that students' beliefs are firmly held and resistant to change. As well as students' existing beliefs influencing perceptions and interpretations of data, some researchers (Fairbrother & Hackling, 1997; Rigano & Ritchie, 1995) report that students deliberately 'fudge' data to obtain what they perceive to be the correct answer.

General Assertion D: Many students do not reject an hypothesis even when their data indicate that this would be a logical conclusion. Instead they modify or attend to their data in such a way that their hypothesis is supported.

In considering ways students address investigations, it appears that there are **aspects of scientific inquiry** that are not explicitly attended to during instruction. They relate to the following; the concept of 'proof' in science; the practice of disconfirming an hypothesis when it is not supported by data; the notion that support for an hypothesis does not necessarily imply support outside the data range; and the practice of ignoring erroneous data. Hence, while these aspects of scientific inquiry remain part of teachers' tacit knowledge they are likely to appear idiosyncratic or inconsistent to students. These aspects of investigating are discussed in more detail.

(a) The concept of **proof** is domain specific. The notion that "Science laws cannot be proved but they can be tested, " (Mackeith, 1989, p. 64) is in conflict with the concept of 'proof' in other domains. Everyday use of the word 'proof' is accepted as "evidence sufficing or helping to establish a fact" (Fowler & Fowler, 1964, p. 980). In other disciplines, evidence leads to proof such as in mathematics and law, appropriate evidence leads respectively to the proof of a theorem and to the proof of guilt.

(b) The notion that it is acceptable to reject an hypothesis. This may appear idiosyncratic to students because it is likely to be contradictory to students' previous experiences at school where correct guesses, predictions and hypotheses are valued and rewarded. In scientific practice, based on Popper's falsification principle (Mackeith, 1989), it is acceptable to propose an hypothesis that may not be supported by data, so long as the data are collected accurately and the processed data and conclusions indicate that the hypothesis is not supported. It needs to be added that learning about the practice of science (procedural knowledge) may not always lead students to accepted understandings of science concepts. Poor experimental work where errors are due to poor equipment, poor control of variables, unrepresentative samples, and to measurement error may result in unreliable data being used to confirm an incorrect hypothesis or reject a correct hypothesis. An example of learning unacceptable views of science occurred with Tammy. She hypothesised "The higher the temperature of the water the more acidic the milk becomes in the presence of lipase," and because her data did not support the hypothesis she later wrote on her questionnaire response that she learned "Temperature doesn't affect the breaking down of fatty acids in the presence of lipase" (Group TN).

(c) Finding support for an hypothesis does not mean that the hypothesis is supported beyond the range of data. For example the hypothesis, "The higher the temperature the slower the rate that lipase will react with milk", is likely to be supported with temperatures approaching 70 to 80 °C. After gathering data which supported this hypothesis, some students claimed support for the following hypothesis, "The lower the temperature the faster the reaction rate." The latter hypothesis is most unlikely to be supported and highlights the need for students to be cautious when extrapolating their findings beyond the range of data.

(d) It is acceptable to ignore some data and not other data. For example during the ukulele investigation and the teacher modelled pitch of a closed pipe investigation students were encouraged to ignore the frequencies of musical notes that sounded different and resulted in an irregular wave pattern. The rationale for this decision was based on the premise that the 'attack' of the musical note had been poorly executed and that the resulting note contained too many irregular harmonics that interfered with the wave analysis. The rejection of data when flaws are identified in the collection process, is an accepted scientific practice. In contrast, it would be erroneous for students to ignore data in the trypsin and lipase investigations that showed that the reaction did not work at high and low temperatures. It is proposed that Year 9 students have difficulty in making appropriate judgements about when it is acceptable to ignore data and when this practice contributes to the rejection of valid and reliable data.

These four aspects of scientific inquiry caused difficulties for some students and were not identified by the teachers or researcher at the start of the instructional program. Consequently, they were not explicitly addressed during the instructional program. The following general assertion addresses these aspects of inquiry.

General Assertion E: Students lack understandings of the conventions and procedures of scientific inquiry, specifically the concept of 'proof' in science, the acceptability of rejecting an hypothesis based on reliable data, the notion that confirming an hypothesis does not imply that the hypothesis is supported beyond the range over which the data were collected, and the notion that data can be ignored when errors are identified in the data gathering process.

Students perceived that they learned more from the first few investigations they performed (Assertion 19) and that they preferred investigations in which they investigated a science problem for which they did not know the answer (Assertion 51). Their teachers said that some students perceived that developing investigation competencies was not learning science and this idea is expressed in Assertion 36. From these assertions two ideas emerge. First, the contribution of novel tasks to learning should not be under emphasised and second, the value that student place on developing conceptual understandings (learning science) needs to be recognised. Thus learning from investigating is more fruitful and purposeful when students engage in new or different tasks from which they develop science understandings that are new to them (General Assertion F). The need to link investigations to scientific understandings is expressed by Coles and Gott (1993) and Hodson (1992). They state, that with out this link the investigations are not science.

General Assertion F: Learning from science investigations is perceived to be more fruitful and purposeful when students engage in novel learning experiences that develop their conceptual understandings of science.

Aspects of the Cognitive Apprenticeship Model of Instruction

Assertions associated with the cognitive apprenticeship model have been grouped as follows; teacher modelling, coaching, scaffolding and fading, articulating, and self-reflective and metacognitive skills (Figure 24). The following discussion contributes to understanding the effectiveness of the cognitive apprenticeship model of instruction as implemented in this study.

Teacher modelled investigations

Assertions 28 and 29 raise concerns about the effectiveness of teacher modelled investigations as a way for students to acquire investigation competency. This is supported by quantitative questionnaire data in which students rated, on average, the teacher modelled investigations fifth (the pitch of a closed pipe investigation) and eighth (the acid and carbonate investigation) out of eight tasks in terms of their contribution to learning about doing investigations. In addition, classroom management became problematic because modelled investigations were difficult to execute in whole-class settings (Assertions 40, 62 and 63). These assertions are collated more generally as follows.

General Assertion G: When learning to perform science investigations, the teacher modelled investigations implemented in this study are not as effective as student investigations.

Assertions

Teacher modelling

28 Students perceive that they learn less about doing investigations from teacher modelled investigations than from student investigations. Q, T, S

29 Students perceive that teacher modelled investigations contribute less than student investigations to their learning about *Planning investigations*, *Conducting investigations*, *Processing data* and *Evaluating investigations*. Q, T, S

30 Students perceive that for the teacher modelled investigations the investigation competency they learn most about is *Conducting investigations*, and for different investigations they learn different sub competencies associated with *Conducting investigations*. Q

40 Teachers perceive that teacher modelled investigations are difficult to implement in whole-class settings because
(a) they are time consuming and the students become off task,
(b) sometimes not all students can observe the data collection, and
(c) setting up the equipment can be fiddly. T, S, V

41 Teacher modelling of investigation competencies may best be implemented with small groups of students so that teacher guidance can be
(a) provided when needed by the group,
(b) highly focused on the difficulties being encountered by the group, and
(c) at a personal level T, S, V

42 The focus of teacher modelled investigations should be on developing a small number of competencies. T, S, V

59 During teacher modelled investigations on-task student behaviour is maintained by requiring students to complete Investigation Planning and Report Sheets. V

60 Different investigations provide opportunities for teachers to model different investigation competencies. V

61 Second hand data can be used effectively to model data processing competencies. V

62 Moving students from their seats so that the data gathering can be observed may result in increased off-task behaviour. V

63 Students appear to lose interest when watching repeated or replicated trials. V

Coaching

27 Some students prefer teachers who give full explanations and some prefer teachers who encourage and assist them to solve their own problems. Q, T

43 Classroom management problems can arise when students find investigations too challenging. T

44 Teachers have different views on how much guidance to give students performing investigations. T, S, V

66 Teachers differ in the type and amount of guidance they give students performing investigations. V

67 Teachers do not always gather appropriate information about students' understandings to make informed decisions about ways to coach students V

Scaffolding and fading

53 Students perceive that the scaffolded Investigation Planning and Report Sheets as implemented in this study are helpful in learning to perform investigations, and that less scaffolding is appropriate at the end of the program. S, V

64 Teachers used the scaffolding provided by Investigation Planning and Report Sheets to organise instruction and to provide a focus for student learning. V

65 Students use the scaffolding provided by Investigation Planning and Report Sheets to ensure that they complete the requirements for investigations. V

Articulating

24 Students perceive that the best way to learn about doing investigations is by doing them with a groups of students. Q, S

25 Most students perceive that they learn more about doing investigations from talking with their peers than from talking with their teacher. Q, S

26 Most students perceive that they learn more about doing investigations from talking with their peers and their teacher early in the instructional program. Q

68 During investigations students engage in discussions about social issues, especially when they encounter difficulties with the investigations. V

69 Group discussions about investigations are mostly concerned with how to complete the task. V

70 Few group discussions appear to have the potential to improve students' understandings of concepts or to develop investigation competencies. V

Self-reflective and metacognitive skills

45 Teachers believe that Year 9 students are not sufficiently mature to reflect on their leaning. T, S

54 Most students do not recognise that self-reflective and metacognitive tasks could be of benefit to their learning. S

Figure 24. Assertions about the cognitive apprenticeship model of instruction

Note: I = Investigation Planning and Report Sheet, S = student interviews, Q = student questionnaires, T = teacher interviews, and V = audio and video data
The code first mentioned indicates the source of the assertion. Subsequent codes indicate triangulation of the data from other sources.

These data contradict the critical role that some researchers (Collins et al. 1989) attribute to teacher modelling in the instructional process. Exponents of the cognitive apprenticeship instructional model would argue that one strength of teacher modelling is that the modelled activity represents authentic practice of the whole target process and that it provides opportunities for situated cognition, thereby reducing the problems of brittle skills and inert knowledge.

Suggested improvements to teacher modelling of science investigations (Assertions 41 and 42) shift the emphasis from modelling the whole investigation at the start of an instructional sequence, to modelling parts of an investigation when needed by groups of students. Hence, the assertions indicate a paradigm shift where modelling becomes one aspect of coaching; of the hints, reminders and focusing of attention on previously unnoticed aspects of the task. Modelling smaller portions of the task also reduces the possibility for information overload. Javela (1996) made the distinction between teacher global modelling in front of the whole-class and situation specific modelling to small groups of students. She claimed the latter had the potential to be more directed at the problems the students were experiencing and to promote more advanced exploratory activity. These notions contrast with Lave's perspective of the function of modelling (in Collins et al., 1989). She believes that through observation a conceptual model of the process is provided and that this contributes to the success of teaching complex skills and obviates the need for lengthy practice of isolated sub-skills. In this study, it may be argued that the conceptual model of the whole task was provided for students by the scaffolded worksheets and that observation of the whole task was not needed to provide a conceptual model. This could also be achieved by providing students with schematic diagrams which show them how the parts of the investigation process form a coherent whole (Mines, 1995; Hackling & Fairbrother, 1996). Hence, only specifically targeted sub competencies need to be modelled, such as measuring the

dependent variable and constructing a table to record data. These ideas are summarised in the following general assertion.

General Assertion H: Teacher modelling should be one of several strategies used to provide guidance for students' learning to perform science investigations; particularly when groups of students need to develop specific competencies.

The data from this study indicate that teacher modelling faired poorly as an instructional strategy and a perusal of different instructional models sheds light on possible reasons. Many instructional models consider activities of the learner to be the focus of instruction and they start with experiences (Renner, 1982), exploration (Karplus, 1977), exposing alternative frameworks (Nussbaum & Novick, 1982) experiential manoeuvres (Erickson, 1979); and orientation and elicitation of ideas (Driver & Oldham, 1986). These behaviours are learner-centred because they describe students' behaviours and/or students' learning experiences. Furthermore, they are perhaps more representative of the constructivist perspective in that they are more focused on the learners' prior understandings, and on identifying and establishing a starting point for instruction. In contrast, the starting point for instruction in the cognitive apprenticeship model is the activity in which the teacher is engaged.

It may be argued that the modelling aspect of the apprenticeship paradigm is better suited to the explication of specific process skills rather than cognitive functions or general procedures from which learners are expected to generalise to other tasks. Hence modelling is likely to be more successful where it is exactly the same as the target process. The medical school adage "see one, do one, teach one" reflects an apprenticeship approach to training doctors. The closeness of a modelled operation with the target process is obvious, and the ensuing "teach one" represents

articulation, and self-reflection and metacognition that is espoused in the cognitive apprenticeship model. Further, a medical student (S. Garnett, personal communication, September 28, 1997) said that observations can be made at two levels of engagement, "You can watch in a detached way or your can think of yourself as doing the operation. For me, I actually put myself in the place of the doctor otherwise I don't get anything out of observing. I suppose its (observing) better than nothing but you really don't know how to do things until you've tried them (operations and procedures) yourself."

Coaching

Teachers' perceptions of their role as coach seemed to depend on the purpose and nature of the activity in which the students were engaged, and their perception of how much help they should give and how much students should be expected to work out for themselves. When performing science investigations students are expected to think about and grapple with ideas before they are provided with hints and advice. In response to students' questions some teachers told students the answers (Mrs Cross, p. 205); some gave indirect responses in that they structured subsequent questions and discussion to lead the students to appropriate solutions (Miss Mills, p. 206); and some expected students to find out answers for themselves and to make their own discoveries and connections (Mr Brogo, p. 168). The latter perspective is consistent with discovery learning that was espoused in the 1970s (Hodson, 1996). These different approaches were expressed in Assertions 44 and 66. As could be expected, students differed in their preferences for the approaches (Assertion 27). These ideas are summarised more generally as follows.

General Assertion I: Teachers differ in the amount and type of coaching they offer to students performing investigations, and students have different preferences regarding teacher coaching.

This finding is supported by research by Sharp and Green (1975). They reported that teachers' beliefs about their teaching role during investigations influenced their practice. For example, those who believed that investigations were about allowing students to discover things for themselves adopted an extreme non-interventionist role acting only as a manager and provider of resources. Mines (1995, p. 14) claims that "the art of skillful questioning appears to be crucial to achieve the balance between giving students suitable guidance and leaving sufficient scope for them to think independently."

Collins et al. (1989, p. 481) stated that teacher observations are the foundation on which decisions about coaching are made; "Coaching consists of observing students while they carry out a task", and it subsequently involves "offering hints, scaffolding, feedback, modeling, reminders, and new tasks aimed at bringing their performance closer to expert performance." While observations may provide a basis for decisions to be made about a students' investigation competencies, it is unlikely that observations alone will provide sufficient information about a student's cognition to be acted on by a coach. More recent studies (Roth, 1995; Javela, 1996) identify verbal interactions as a source of information for coaching, and asking students questions is a quick way to get information and likely to be more effective. While the concept of providing feedback and guidance on students' performance is addressed in the instructional model, the way teachers obtain information on which to base feedback and guidance is sketchy and still evolving. Solely using observations as the basis for making coaching decisions during science investigation lessons is likely to yield incomplete, superficial information even when students are observed for relatively long periods of time. Richer data may be gathered when teachers use their knowledge of students' understandings of science concepts and investigation competencies, together with their knowledge about the nature of the task, to question students

about investigations. From students' responses to these questions teachers will then be able to coach students in the zone of proximal development (Vygotsky, 1986).

Scaffolding and Fading

Students perceived that the scaffolded IPRSs used in the study were helpful in learning to perform investigations (Assertion 53) and they used them to confirm that they had completed the requirements for the investigation (Assertion 65). The IPRSs provided students with a conceptual framework for the investigations. The provision of a conceptual model is important because it provides the learner with an advanced organiser, a guide for practice, and a structure for making sense of the feedback (Collins et al. 1989).

Teachers used the scaffolded IPRSs to organise instruction and to provide a focus for students' learning (Assertion 64). Also, they were used to organise instruction during the teacher modelled investigations, and to structure feedback after the students had completed the investigations. With the teacher modelled investigations, Mrs Cross used the IPRSs to keep students on task (Assertion 59). Therefore, in terms of the cognitive apprenticeship model of instruction this research has recognised that scaffolded worksheets can be useful for teachers as well as for students. These ideas are expressed as follows.

General Assertion J: Scaffolded Investigation Planning and Report Sheets provide a conceptual model for investigations and are useful for teaching and learning, assessment, and classroom management.

As implemented in this study, scaffolding was a strength of the instructional model. The interpretation of scaffolding in this research differed slightly from other studies. In this study it was defined as "A predetermined strategy or structure used

to facilitate learning that was based on the teacher's conceptual model of the task." In other studies (Collins, et al. 1989; Hennessy, 1993) scaffolding applies to an impromptu structure that the teacher applies in response to the needs of the learner. "Subsequent interpretations and applications of the notion of apprenticeship have without exception focused on the tutor's implicit theory of the learner as being a crucial element of the scaffolding process" (Hennessy, 1993, p. 12). Therefore the teacher must display sensitivity to the "learner's current needs, knowledge structure and performance characteristics" and this needs to interact with the "tutor's theory of the task or problem" (p. 12). These developments of the model acknowledge that the needs and characteristics of the learner play a central role in informing scaffolded instruction.

Articulating

An inconsistency in the data was the different perceptions that students and the researcher formed about students' learning from discussions with other students. Most students perceived that they learned more about doing investigations from talking with their peers than from talking with their teacher (Assertion 25) and this may have influenced their belief that the best way to learn about doing investigations was in a group setting (Assertion 24). The audio and video tapes captured students engaging in social discourse, particularly when they encountered difficulties with the investigation (Assertion 68). Also, students' discussions relating to the investigation centred on how they would accomplish the task (Assertion 69) as this was likely to be their primary goal. Discussions that had the potential to improve conceptual understandings or to develop investigation competencies were rarely observed (Assertion 70) as they were likely to be secondary goals for students. These ideas are expressed as follows.

General Assertion K: Students' on-task discussion is mainly associated with how to complete the investigation and students perceive they learn more about doing investigations from talking with their peers than their teacher.

Studies of school science laboratories (Christensen & McRobbie, 1994; Gallagher & Tobin, 1987; Kempa & Ayob, 1995; Tobin, 1990; Roychoudhury & Roth, 1996) report varied findings about students' learning from group work and, by inference, learning from articulation amongst students. The findings by Christensen and McRobbie, have parallels with this study because they similarly report that the researchers and students had different perceptions of the learning that occurred. They describe laboratory lessons in which on-task behaviours focused on the procedures to complete the task, however, despite this "students believed quite strongly that practical work was vital to their understanding of concepts" (p. 58). Roychoudhury and Roth, and Kempa and Ayob report more positive outcomes. Roychoudhury and Roth say that by far the majority of students viewed group work positively, and that most perceived that there were benefits in pooling ideas. These findings are also consistent with students' perceptions of group work in this study. Roychoudhury and Roth, and Linn and Burbules (1993) both report that group learning is not liked by all students and again this is consistent with this study because one of the students, Olive, stated that she preferred to do things by herself rather than by discussing matters with the group because the group had difficulties reaching a consensus (p. 187). From pencil and paper tests Kempa and Ayob report that there was a satisfactory level of achievement from group work and that a significant amount of learning occurred from other students through shared knowledge. In this study, although large gains were recorded in the pencil and paper TOSIS, and in students' investigation competencies as measured by investigations (IPRS) it is not possible to apportion contributions from different aspects of the instructional model.

The notion of collaborative group work is not a feature of the cognitive apprenticeship instructional model, although it is a feature of instruction in science laboratory lessons. The articulation described in the instructional model represents an expert-novice perspective where it is scaffolded and constructed by the expert to bring the novice's behaviour and skills closer to that of the expert. In contrast, within classroom settings much of the articulation is between students. Consequently, it may be argued that applying the cognitive apprenticeship model in situations where expert-novice interactions and articulation are but one form of interaction is an extension or modification of the original model.

Self-Reflective and Metacognitive Skills

Students participated in three metacognitive exercises; the student questionnaires, the last section of the IPRSs, and the worksheet on marking an investigation. In addition, Class SC assessed their own investigations. The data from teacher and student interviews indicate that teachers believe students were not sufficiently mature to reflect on their learning (Assertion 45) and students were reserved in their perceptions of the benefits from this activity (Assertion 54). Collectively these ideas are expressed as follows.

General Assertion L: Most students have poor metacognitive skills and do not recognise the value of these skills.

Collins et al. (1989) advocated two strategies to promote reflection; the comparison of expert and novice performances on problem solving processes, and students' self-analysis of the process. In this study both strategies were applied. Opportunities for the former were created when students' investigations were assessed either by the teachers (Class TN and Class TC) or by the students (Class SC). In both instances students compared their investigation performances with

their teacher's expectations; either during teacher feedback to the whole-class (Classes TN and TC) or against a prepared master answer sheet (Class SC). Hence students matched their performance with that of an expert. Self-analysis was achieved when students completed the last section of the IPRS (Appendix F) in which they were asked to reflect on their performance, and when they completed the questionnaires and were asked about their learning. Although these opportunities were afforded to students it appeared that they could see no purpose in, or benefit from them. Clearly students need to perceive that self-reflective and metacognitive practices lead to improved learning before they are prepared to engage in these practices.

Assessment

The study involved different assessment regimes for the three participating classes. For Class TN students' investigations were teacher assessed and norm referenced; Class TC investigations were teacher assessed and criterion referenced and Class SC investigations were student assessed and criterion referenced. The statistical analysis, two-way ANOVA, of the pencil and paper TOSIS that measures a total score for investigation skills and concept areas, indicated that there was no significant difference between the performances of the classes on the pretest or on the posttest. There was, however, an effect for test occasion because the classes improved their performances from the pre to posttest, $F(1, 2) = 66.68$ $p < .01$, and no interaction effects. These findings are expressed as General Assertion M.

General Assertion M: For students experiencing teacher assessed norm referenced assessment, teacher assessed criterion referenced assessment and student assessed criterion referenced assessment there is no significant difference in investigation competency as measured by the students' total score on the pencil and paper Test of Science Investigation Skills.

The descriptive statistics for the total TOSIS score (Table 1 and Figure 6) showed slightly more modest gains for Class SC than for the other classes. Similarly the descriptive statistics for the following specific skill and concept areas indicated that the gains for Class SC were more modest than for Classes TN and TC; *Planning an investigation, Drawing conclusions, Identifying methodological limitations*, and *Understanding the concepts of hypothesis, theory, data and conclusions* (Table 2 and Figures 9 to 12).

Assertions

Teacher assessed norm referenced

46 Teachers found it difficult to rank students' Investigation Planning and Report Sheets in order of achievement. [T]

47 Teachers found norm referenced assessment of the investigations to be very time consuming. [T]

71 The implementations of teacher assessed norm referenced assessment and teacher assessed criterion referenced assessment were very similar. [V]

Teacher assessed criterion referenced

33 Students do not acknowledge that criterion referenced aspects of assessment are helpful in providing feedback on investigation performance. [Q]

48 Teachers found criterion referenced assessment of the investigations to be very time consuming. [T]

49 Teachers found that with criterion referenced assessment it was easier to assign the grades of individual students group by group. [T]

56 Some students believe the teacher assessed criterion referenced assessment as implemented in this study is helpful because it highlights investigation competencies that could be improved. [S]

Student assessed criterion referenced

32 Students who assess their own investigations perceive that they learn less from this process than students who have their investigations assessed by teachers. [Q]

35 Students who have not assessed their own work are less likely to see the value in this assessment procedure than students who have. [Q]

50 Teachers do not like student self-assessment because they do not get feedback on students' performances and as a consequence they find it difficult to address students' errors. [T, S]

57 Students are uncertain about judging their investigation performance based on predetermined criteria. [S]

58 Students' self-assessments of their investigations are superficial and they offer reduced opportunities for teacher feedback.[S, V]

74 Some students using student assessed criterion referenced assessment, assume their responses to questions are incorrect because their language differs from that on master answer sheets. [V]

75 Some students using student assessed criterion referenced assessment are unable to access information on master answer sheets because they lack adequate science background knowledge. [V]

76 Some students using student assessed criterion referenced assessment fail to identify all the relevant feedback on master answer sheets. [V]

Other assertions

31 Students perceive that the worst way to learn about doing investigations is by correcting or marking investigations. [Q]

34 Students' perceptions of the amount of learning resulting from teacher and student assessment vary widely. [Q]

55 *Some students believe whole-class feedback on investigation performance is helpful in improving performance.* [S]

72 *When receiving verbal teacher feedback about investigation performance, most students are inattentive in whole-class settings.* [V]

73 *Following teacher feedback about investigation performance, most students do not make written corrections to their investigation Planning and Report Sheets.* [V]

Figure 25. Assertions about the three assessment regimes

Note: 1. Assertions in italics contain contradictory elements and are deleted from further discussion.
2. I = Investigation Planning and Report Sheet, S = student interviews, Q = student questionnaires, T = teacher interviews, and V = audio and video data
The code first mentioned indicates the source of the assertion. Subsequent codes indicate triangulation of the data from other sources.
3. Assertions presented in italics are not supported by the triangulation of data.

The TOSIS and IPRS measured different aspects of investigating. The TOSIS provides a total test score and also scores on specific skills and concept areas from a pencil and paper test. The IPRS provides more holistic and authentic assessments on a practical task. Therefore, although these measures are complementary it is reasonable to expect that some differences may emerge.

The classes' pretest and posttest performances on an investigation, as indicated by the IPRSs, were analysed by the Kruskal-Wallis One-Way Analysis of Variance. There was no significant difference between the classes before the instructional program ($p < .01$), however, following the instructional program there was a significant difference in the classes' performance for some skills; for the planning phase ($H = 9.76$, $df = 2$, $p < .01$) and also for the conducting phase

(H = 19.36, df = 2, p < .01). Descriptive statistics from the IPRS (Tables 3 & 4) and from the TOSIS (Table 2 and Figures 9 - 12) suggest Class SC's performance was lower than the other two classes in planning and conducting investigations. This findings is expressed as General Assertion N.

General Assertion N: Students participating in self-assessed criterion referenced assessment of science investigations made more modest improvements than students who had their investigations assessed by their teachers for *Planning* and *Conducting investigations* as measured by performance on the Investigation Planning and Report Sheet; and for *Planning an investigation, Drawing conclusions, Identifying methodological limitations* and *Understanding the concepts of hypothesis, theory, data and conclusions* as measured by the pencil and paper Test of Science Investigation Skills.

The qualitative data previously interpreted and summarised as assertions, are clustered under the headings; teacher assessed norm referenced, teacher assessed criterion referenced, student assessed criterion referenced, and general assertions about assessment in Figure 25. The triangulation of these data indicates that some assertions are contradictory and these are presented in italics. For example, Assertion 55: Some students believe that whole-class feedback on investigation performance is helpful in improving performance, appears inconsistent with the observations made from the audio and video data and expressed as Assertion 72: When receiving verbal teacher feedback about investigation performance, most students are inattentive in whole-class settings. In addition, the observation that most did not make corrections to their IPRS when they were warranted (Assertion 73) raises doubts about the helpfulness of whole-class feedback in improving performance as was postulated in Assertion 55: Some students believe whole-class feedback on investigation performance is helpful in improving performance. Similarly, Assertion 56: Some students believe the teacher assessed criterion

referenced assessment as implemented in this study is helpful because it highlights investigation competencies that could be improved, is inconsistent with Assertion 33: Students do not acknowledge that criterion referenced aspects of assessment are helpful in providing feedback on investigation performance. As a consequence, consideration needs to be given to the nature of the data sources. Assertions 55 and 56, that present positive perspectives, were formulated from student interview data. Assertions 71 and 72 which were generated from the audio and video data, and Assertion 33 from student questionnaires, are less positive. Although is not possible to establish which of the data sources has the greatest credibility, one explanation postulated with caution is that the students who were interviewed were more positive in their responses because they may have wished to please the researcher. Another explanation is that it is difficult to interpret the amount and quality of learning from observational data. Assertions which appeared to be contradictory are omitted from the following discussions.

The assertions about teacher assessed norm referenced and teacher assessed criterion referenced assessment are similar and hence support quantitative data that indicates that there is no difference in the learning outcomes for Class TN and Class TC. Both assessment procedures were time consuming (Assertions 47 and 48) and more importantly, the assessment feedback procedures as described in detail from the audio and video data were similarly implemented (Assertion 71). Therefore it may be postulated that in terms of improving investigation competencies, teachers' assessment procedures including norm referenced assessment and criterion referenced assessment, did not differentially impact on students' achievement of investigation competency. These ideas are expressed as follows.

General Assertion O: The development of students' investigation competencies is not influenced differentially by teacher norm referenced and teacher criterion referenced assessment as implemented in this study.

In contrast with teacher assessment, students who assessed their own investigations, Class SC, made more modest gains than Classes TN and TC for *Planning investigations* and *Conducting investigations*, as indicated by the pre and posttest IPRS. In examining the assertions, however, is comes as no surprise that Class SC did not perform as well as the other classes on some aspects of investigating. On average they rated their learning less positively than students whose work was teacher assessed (Assertion 32). Although some students in Class SC perceived that they were able to look more critically at their work, this was not an overall trend in the data. Evidence of an "empowering impact" on students was not observed as it was in the Klenowski study (1995, p. 20). Indeed, shortcomings in the student assessment procedure were identified and it is possible that the potential to improve metacognitive skills was offset by the lack of opportunity that students had for high quality feedback from their teachers (Assertion 58). The lack of teacher guidance afforded by this type of assessment is evident from Assertions 74, 75 and 76. Respectively, these assertions refer to students assuming that their responses were incorrect when in fact they were correct; students' inabilities to comprehend teacher responses on the master answer sheet; and to identify all the relevant feedback from the master answer sheet. The corollary to these assertions is that teachers recognised that they did not receive sufficient feedback on the students' performance and, as a result, they found it difficult to address students' errors (Assertion 50). It is also likely that students perceived that the teachers were less interested in their performance because the teachers did not take the time to assess their work. As a result, students' motivation to learn may have been less than for students in other classes. These ideas are summarised in the following general assertion.

General Assertion P: Students who participate in criterion referenced self-assessment of investigations lack opportunities for high quality teacher feedback. In addition, teachers obtain less feedback on students' performance and as a consequence the potential for teacher guidance is affected.

Summary of the Chapter

This Chapter collated and discussed data gathered from multiple data sources; from the pretests and posttests (the Test of Science Investigation Skills and the Investigation Planning and Report Sheets); student questionnaires; teacher and student group interviews, and from audio and video recordings. Assertions generated from these data sources were presented about the themes of the research, investigation competencies (Figure 23), the cognitive apprenticeship model of instruction (Figure 24) and assessment (Figure 25). From these data 16 general assertions were formulated. Schematically the generation of these general assertions is presented in Figure 22 and they are presented in Figure 26.

General Assertions

Investigation competencies

A When students do not have an appropriate conceptual model of science investigations they lack direction, they do not pay attention to relevant details and/or they often do not complete the investigation.

B Students' abilities to perform investigations are influenced by their understanding of the science content knowledge that underpins the investigation, and their knowledge and understanding of investigation skills and procedures.

C The difficulty of an investigation is influenced by factors associated with the nature of the task such as the complexity of setting up the equipment, the type of variables (whether they are discrete or continuous) and way the variables are measured and controlled.

D Many students do not reject an hypothesis even when their data indicate that this would be a logical conclusion. Instead they modify or attend to their data in such a way that their hypothesis is supported.

E Students lack understandings of the conventions and procedures of scientific inquiry, specifically the concept of 'proof' in science, the acceptability of rejecting an hypothesis based on reliable data, the notion that confirming an hypothesis does not imply that the hypothesis is supported beyond the range over which the data were collected, and the notion that data can be ignored when errors are identified in the data gathering process.

F Learning from science investigations is perceived to be more fruitful and purposeful when students engage in novel learning experiences that develop their conceptual understandings of science.

The cognitive apprenticeship model of instruction

G When learning to perform science investigations, the teacher modelled investigations implemented in this study are not as effective as student investigations.

H Teacher modelling should be one of several strategies used to provide guidance for students' learning to perform science investigations; particularly when groups of students need to develop specific competencies.

I Teachers differ in the amount and type of coaching they offer to students performing investigations, and students have different preferences regarding teacher coaching.

J Scaffolded Investigation Planning and Report Sheets provide a conceptual model for investigations and are useful for teaching and learning, assessment, and classroom management.

K Students' on-task discussion is mainly associated with how to complete the investigation and students perceive they learn more about doing investigations from talking with their peers than their teacher.

L Most students have poor metacognitive skills and do not recognise the value of these skills.

Assessment

M For students experiencing teacher assessed norm referenced assessment, teacher assessed criterion referenced assessment and student assessed criterion referenced assessment there is no significant difference in investigation competency as measured by the students' total score on the pencil and paper Test of Science Investigation Skills.

N Students participating in self-assessed criterion referenced assessment of science investigations made more modest improvements than students who had their investigations assessed by their teachers for *Planning* and *Conducting investigations* as measured by performance on the Investigation Planning and Report Sheet; and for *Planning an investigation, Drawing conclusions, Identifying methodological limitations* and *Understanding the concepts of hypothesis, theory, data and conclusions* as measured by the pencil and paper Test of Science Investigation Skills.

O The development of students' investigation competencies is not influenced differentially by teacher norm referenced and teacher criterion referenced assessment as implemented in this study.

P Students who participate in criterion referenced self-assessment of investigations lack opportunities for high quality teacher feedback. In addition, teachers obtain less feedback on students' performance and as a consequence the potential for teacher guidance is affected.

Figure 26. General assertions formulated from the research

CHAPTER 10

CONCLUSIONS

Overview of the Chapter

This Chapter presents a summary of the research in which the major findings are presented and implications for teaching and the theoretical framework are addressed. The limitations of the research are discussed, recommendations for future research are made and the contribution of the research to teaching and learning is outlined.

Summary and Findings of the Research

The purpose of this study was to develop, implement and evaluate a Year 9 science laboratory investigations program which included open investigations and was based on the cognitive apprenticeship model of instruction and linked to three assessment procedures. The study was conducted at an all girls school in Perth, Western Australia over a seven month period of time. Three classes of students (n = 66) with similar academic abilities were selected from the Year 9 cohort. Students participated in three, 10 lesson segments, comprising two worksheets, two teacher modelled investigations and six investigations. The investigations involved students, working in groups of three, conducting laboratory work to determine relationships between variables and students chose the independent variable they would examine. The students planned and conducted their own investigation, processed the data and evaluated the findings of the investigation. The three classes had their investigations assessed differently; teacher assessed norm referenced, teacher assessed criterion referenced and student assessed criterion referenced.

Different data sources were used to gain multiple perspectives of the learning milieu. Students completed the pencil and paper Test of Science Investigation Skills (TOSIS) as a pretest and posttest, and they also performed similar pretest and posttest investigations which involved the completion of Investigation Planning and Report Sheets (IPRSs) which were assessed using Student Outcome Statements (Education Department of Western Australia, 1997). Students' investigation competencies were determined from these tests. Students also completed questionnaires after each 10 lesson sequence. Groups of three students were interviewed after the 10 lesson sequences and these students were also audio and video recorded as they performed investigations. The teachers were interviewed after each lesson sequence, and they were audio and video recorded as they conducted the modelled investigation and when feedback was provided after the investigations of Classes TN and TC had been assessed.

The following research questions were addressed and the findings are presented.

Research Question 1

What science investigation competencies and understandings are developed by students during the instructional program implemented in the study and what difficulties do students experience?

Statistically significant improvements were recorded in students' knowledge and understandings of skills and concept areas as measured by the total TOSIS score. A two-way Analysis of Variance (ANOVA) indicated an effect for pre and post test occasion of $F(1, 2) = 66.68$, $p < .01$. Descriptive statistics indicated the development of the following specific skill and concept areas; *Identifying variables, Writing an hypothesis, Planning an investigation, Drawing conclusions, Identifying*

methodological limitations, and *Understanding the concepts of hypothesis, theory, data and conclusions*. Complementary data gathering procedures based on students' investigation performances and assessed using IPRSs indicated that students improved at *Planning investigations, Conducting investigations, Processing data* and *Evaluating investigations*. The Wilcoxon Matched-Pairs Signed-Ranks test was used to compare the pre and posttest levels of performance for the classes and for all phases of investigating the posttest levels were significantly higher than the pretest levels, at the $p < 0.01$ level for one tailed tests. In addition, student questionnaires and student group interview data indicated that students perceived that they develop not only these skills but also the skills of *Working cooperatively, Attending to detail, Managing time* and *Being organised*.

Six general assertions (A to F) were generated about students' investigation competencies from 29 specific assertions and they primarily focused on difficulties in learning to perform investigations.

A When students do not have an appropriate conceptual model of science investigations they lack direction, they do not pay attention to relevant details and/or they often do not complete the investigation.

B Students' abilities to perform investigations are influenced by their understanding of the science content knowledge that underpins the investigation, and their knowledge and understanding of investigation skills and procedures.

C The difficulty of an investigation is influenced by factors associated with the nature of the task such as the complexity of setting up the equipment, the type of variables (whether they are discrete or continuous) and way the variables are measured and controlled.

D Many students do not reject an hypothesis even when their data indicate that this would be a logical conclusion. Instead they modify or attend to their data in such a way that their hypothesis is supported.

E Students lack understandings of the conventions and procedures of scientific inquiry, specifically the concept of 'proof' in science, the acceptability of rejecting an hypothesis based on reliable data, the notion that confirming an hypothesis does not imply that the hypothesis is supported beyond the range overwhich the data were collected, and the notion that data can be ignored when errors are identified in the data gathering.

F Learning from science investigations is perceived to be more fruitful and purposeful when students engage in novel learning experiences that develop their conceptual understandings of science.

Research Question 2

In the teaching and learning of science investigation competencies how effective is the cognitive apprenticeship model of instruction?

The cognitive apprenticeship model of instruction is effective in the teaching and learning of science investigation competencies. Improvements in students' investigation competencies were addressed by Research Question 1.

The instructional model as implemented in the study, comprised modelling, coaching, scaffolding and fading, articulating, and self-reflection and metacognition. A holistic approach to the teaching and learning program was adopted and these teaching strategies were not implemented in isolation but mutually shaped the instructional program.

The teacher **modelling** of two science investigations was perceived to be less effective than other instructional strategies for learning investigation competencies. Classroom management issues affected the success of the modelled investigations and students indicated that they would prefer to do the investigations in a group and obtain help from the teacher when needed. Based on student and teacher opinion it was proposed that teacher modelling could be more effective if it were implemented as a coaching strategy to illustrate specific aspects of an investigation. Two general assertions summarise these findings.

G Teacher modelled investigations, as implemented in this study, are not as effective as other teaching strategies for learning about performing science investigations.

H Teacher modelling should be one of several strategies used to provide guidance for students' learning to perform science investigations; particularly when groups of students need to develop specific competencies.

Teacher **coaching** comprised helping students to perform investigations. Class SC experienced teacher coaching during the investigations but did not receive coaching after the investigations. Classes TC and TN received coaching during and after the investigations, and following the assessment of their investigations teachers spent approximately 20 minutes addressing errors that students had made on the investigation. Although all classes significantly improved at performing investigations as demonstrated by the Wilcoxon Matched-Pairs Signed-Ranks test, the results of a Kruskal-Wallis One-Way Analysis of Variance indicated that the classes did not all perform equally well on the posttest for *Planning investigations* ($H = 9.79$, $df = 2$, $p < .01$) and *Conducting investigations* ($H = 19.36$, $df = 2$, $p < .01$). The descriptive statistics from the TOSIS and the IPRSs indicate that Class SC did not perform as well as Classes TN and TC on some skills and

competencies. These findings suggest that teacher coaching received during and after the investigations had a more positive influence on students' investigation competency than coaching only received during the investigations. These interpretations are proposed with caution because, although the classes were of similar abilities, the study unfolded in a naturalistic setting without control of interferring variables.

The nature of the assistance that teachers offered students appeared to be dependent on the teacher's perception of their role during instruction and learning. This is expressed as follows.

I Teachers differ in the amount and type of coaching they offer to students performing investigations, and students have different preferences regarding teacher coaching.

Scaffolding and fading were perceived as effective instructional strategies, when implemented in this study. Scaffolding was interpreted as a predetermined strategy or structure based on the teacher's conceptual model of the task and used to shape and facilitate learning. Scaffolded Investigation Planning and Report Sheets (IPRSs) were used to provide a conceptual model for the investigations. Fading was the withdrawal of scaflolding as students became more familiar and increasingly competent at investigating. Teachers found the scaffolded IPRSs useful for providing a focus for student learning and for organising instruction. Students found the scaffolded IPRS useful in assisting their learning and in ensuring that they had met the requirements of the task. These findings have been summarised a follows.

J Scaffolded Investigation Planning and Report Sheets provide a conceptual model for investigations and are useful for teaching and learning, assessment, and classroom management.

Articulation occurred between students, and between teachers and students. The cognitive apprenticeship model does not address student-student articulation although it is a feature of collaborative group work in science laboratory classes. In this study the effectiveness of learning from other students was not explicitly researched, however, students perceived that the best way to learn about doing science investigations was by working in a group of students. In contrast, the audio and video data revealed that students' on-task articulation focused on how they would complete the task, and few discussions that could contribute to the improvement of students' understandings of science concepts or the development of investigation competencies were observed. Nonetheless significant gains were recorded in students' investigation competencies. Teacher-student interactions are representative of an expert-novice learning milieu that characterises the notion of apprenticeship and are consistent with the notion of coaching. The following general assertion summarises the findings about articulating.

K Students' on-task discussion is mainly associated with how to complete the investigation and students perceive they learn more about doing investigations from talking with their peers than their teacher.

Self-reflection and metacognitive skills were not perceived to be of value by the students, and teachers thought that students' skills were poor. Although the instructional model documents the importance of these skills it does not address ways of helping students to appreciate their importance .

L Most Year 9 students have poor metacognitive skills and do not recognise the value of these skills.

Research Question 3

> **What effect do different assessment procedures including teacher assessed norm referenced, teacher assessed criterion referenced and student assessed criterion referenced assessments have on students' learning of investigation competencies?**

The investigations that the classes completed were assessed differently; Class TN was teacher assessed and norm referenced, Class TC was teacher assessed and criterion referenced and Class SC was student assessed and criterion referenced. Students' total scores on the pencil and paper Test of Science Investigation Skills (TOSIS) were used to measure achievement of specific skills and concept areas associated with investigating. A two-way Analysis of Variance indicated that on the pre and on the posttest there was no significant difference for class.

M For students experiencing teacher assessed norm referenced assessment, teacher assessed criterion referenced assessment and student assessed criterion referenced assessment there is no significant difference in investigation competency as measured by the students' total score on the pencil and paper Test of Science Investigation Skills.

The descriptive statistics for specific skills and concept areas measured by the TOSIS (Table 2 and Figures 9 to 12) showed that Class SC achieved smaller gains than Classes TN and TC for *Planning an investigation, Drawing conclusions, Identifying methodological limitations,* and *Understanding the concepts of hypothesis, theory, data and conclusions.* These data are supported by analyses of pre and posttest performances on the IPRS. Although the Kruskal-Wallis One-Way Analysis of Variance indicated that there was no significant difference between the classes before the instructional program on the IPRS, at the $p < 0.01$ level, there was a significant difference between the classes on the posttest for *Planning*

Investigations (H = 9.76, df = 2, p < .01) and *Conducting Investigations* (H = 19.36, df = 2, p < .01). Fewer students in Class SC attained the benchmark of Level 5 performance for *Planning investigations* and *Conducting investigations* on the posttest than students in Classes TN and TC (Table 4). The TOSIS data also indicates that Class SC scored more modest gains than the other classes for some skill and concept areas. These data are summarised as follows.

N Students participating in self-assessed criterion referenced assessment of science investigations made more modest improvements than students who had their investigations assessed by their teachers for *Planning* and *Conducting investigations* as measured by performance on an Investigation Planning and Report Sheet; and for *Planning an investigation, Drawing conclusions, Identifying methodological limitations* and *Understanding the concepts of hypothesis, theory, data and conclusions* as measured by the pencil and paper Test of Science Investigation Skills.

Clearly the data are equivocal in that the total TOSIS score indicated that there was no significant difference between the classes, while some TOSIS subtest scores and the IPRS results indicated that Class SC made more modest improvements than Classes TN and TC on some investigation competencies. Qualitative data based on student questionnaires, teacher and student group interviews, and audio and video data augmented these findings and it was observed that Class SC lacked opportunities for high quality feedback from their teachers. The following general assertions were postulated.

O The development of students' investigation competencies is not influenced differentially by teacher norm referenced and teacher criterion referenced assessment as implemented in this study.

P Students who participate in criterion referenced self-assessment of investigations lack opportunities for high quality teacher feedback. In addition, teachers obtain less feedback on students' performance and as a consequence the potential for teacher guidance is affected.

Implications for Teaching

The implications for teaching, curriculum development and science education arising from this research are associated with developing understandings of the factors contributing to investigation competency and understandings of difficulties that students experience when learning to investigate, the strengths and weaknesses of the cognitive apprenticeship instructional model, and different assessment regimes.

An understanding of factors that contribute to science investigation competency is needed by teachers, curriculum developers and science educators so that they are able to plan, select and/or design learning experiences that develop the range of skills and competencies needed for successful scientific inquiry. For example, factors which affect the difficulty of investigations need to be considered, and investigations need to be selected with care so that students work in their zone of proximal development (Vygotsky, 1986). Investigations should be linked to the development of science concepts because students then regard them as being more purposeful. Teachers, science educators and curriculum developers need to be aware of the influence and persistence of students' prior beliefs in shaping the ways that students interpret data, and aware of students' confirmatory bias because they modify and/or interpret data so that it confirms existing beliefs.

Teachers should use scaffolded and guided learning experiences to improve investigation performance. Scaffolded Investigation Planning and Report Sheets provide students with a conceptual model of investigations and are useful for structuring teaching and learning, assessment and classroom management. Teachers need to consider that modelling may be more effective with small groups when it is needed by the group, highly focused and at a personal level. Concept areas that need addressing during instruction include aspects of scientific inquiry such as the concept of scientific proof; the notion that it is acceptable to reject an hypothesis; the notion that supporting an hypothesis does not mean that the hypothesis is supported beyond the data range; and that it is acceptable to ignore some data and not other data. Students should be provided with opportunities to engage in self-reflective and metacognitive practices and the benefits of developing these skills need to be realised through practice and instruction.

The findings indicate that teachers should play a more active role in informal formative assessment during science investigations. The effectiveness of current informal formative assessment practice needs to be examined because students perceive they learn more about doing investigations from talking with their peers than from their teachers, and because when teachers do not formally assess students' investigations (Class SC) students lack quality feedback on their performances. During investigations teachers need to explore ways of obtaining information about students' understandings of the investigation. It is unlikely that observations of group work are sufficient to yield appropriate information on which teachers can base coaching decisions. Therefore teachers need to ask students questions as they perform investigations so that appropriate guidance may be given.

Implications for the Theoretical Framework

The study was based on the contructivist perspective of learning and the instructional methodology was grounded on the notion of cognitive apprenticeship.

Learning was evident from a comparison of pretest and posttest investigation competencies. Notions about learners' constructing their own understandings emerged from a variety of students' responses to the IPRSs, particularly when students' beliefs were inconsistent with established scientific theory. The impact of students' prior beliefs on shaping subsequent learning experiences was evident when they modified or attended to their data so that existing beliefs were confirmed, instead of acknowledging and attending to data that could challenge these beliefs. This supports previous findings about the conviction with which people hold established beliefs (Kuhn, 1992; Novak, 1988; Nussbaum & Novick, 1982).

The cognitive apprenticeship instructional model represents an expert-novice perspective of learning in which the novice engages in authentic practice. In this research, investigating represents the practice of scientific research. The teacher modelling aspect of the instructional model may be better suited to learning skills in a tightly defined domain where a high expert-novice ratio is possible. In such situations the expert may be able to determine more easily the novices' prior knowledge and scaffold instruction appropriately. The strength of the instructional model, as implemented in this study, was in the notion of providing guidance for learning, in the form of scaffolding and coaching.

Limitations of the Research

Several limitations of the study are identified. These are associated with the data gathering, data analysis and interpretation, and generalisability of the findings.

Limitations of the Data Gathering

Sampling limitations resulted from the nature of the research and the setting in which the research was conducted. Approximately 66 middle ability students, comprising three classes from the Year 9 cohort at an all girls' school participated in the study. The consequences of this were that data were not obtained for high and low ability students and that data related to a small number of girls from a high socio-economic group. Four teachers implemented the cognitive apprenticeship model of instruction in three classes so that teachers' perceptions and beliefs were based on a small sample of well qualified and experienced science teachers with both science and education qualifications. Although students completed the questionnaires and answered questions about the instructional model, their perceptions were grounded on the work of one or two teachers who implemented the instructional approach within their class. Only one class participated in each of the three assessment regimes, teacher assessed norm referenced assessment, teacher assessed criterion referenced assessment and student assessed criterion referenced assessment.

The study was naturalistic in that it unfolded in a real world setting. Because of this, the learning milieu was not controlled and different influences may have impacted on students' achievement of investigation competencies. For example, other programs of study in the school may also have contributed to students' improved investigation competencies, and in addition, improvements may have been due to maturation factors because the research took place over a seven month period. Different teachers implemented the programs in the three classes and although there was some rotation of teachers amongst classes, teachers may have affected the achievement of the classes differently. The classes, however, were selected on the basis of Year 8 test results and structured to be of similar academic ability.

With the performance assessment of students' IPRS, it is recognised that the assessment is heavily dependent on the nature of the tasks and on antecedent instruction (Gipps, 1994). Numerous researchers (Haertel, 1993; Linn, Baker & Dunbar, 1991; Shavelson & Baxter, 1992) state that even in tightly constrained situations in which parallel tasks are kept similar, it is difficult to make two tasks function the same way. In this research, parallels were drawn between the pretest and posttest investigations, Chapter 3. There was, however, a general perception amongst the teachers and the researcher that the trypsin posttest investigation was easier than the lipase pretest investigation because students found the dependent variable easier to measure. As a consequence, pre to post test gains in students' investigation competencies would have been greater than if a more difficult posttest investigation had been used.

Limitations of the Data Analysis and Interpretation

A range of data gathering procedures was used including pre and posttests, observation of audio and video data, student and teacher interviews and students' questionnaires. From these data, assertions and subsequently general assertions/findings were made. The generation of assertions required a search for patterns and links in the data. As a consequence, the process dealt with frequent events well, and addressed infrequent events less well because emergent patterns were not evident (Erickson, 1986).

The data analysis was conducted several months after the data collection. This resulted in missed opportunities for the researcher to use the analysis to shape and focus subsequent student and teacher interview questions in order to extract richer information. As a result, reasons and explanations for students' perceptions emerging in Assertions 28-30 and 36 have not been developed.

The research questions were wide ranging in focus and addressed students' investigation competencies, the cognitive apprenticeship model of instruction and assessment procedures. This holistic approach to the teaching of and learning about science investigations, mitigates against highly detailed and narrowly defined data. In tying the threads of the research together, fine grained data tends to be overlooked in the search for big issues and findings, and the findings themselves are at risk of becoming so general that they lack utility.

Generalisability of the findings

The degree to which the findings of this study can be generalised to other settings depends on the extent to which the audience can identify with the environment in which the research was conducted.

Recommendations for Future Research

The study was implemented using an holistic approach to the teaching and learning of science investigation competencies and the data were numerous in source and wide ranging in scope. While similar studies would improve the generalisability of the findings more focused research could address specific issues such as:

(a) the effectiveness of different instructional models on students' learning of investigation competencies, particularly models which have as their starting point activities in which the student is more actively engaged in exploration (Karplus, 1977), or in orientation and elicitation of ideas (Driver & Oldham, 1986);

(b) ways to develop students' abilities to consider evidence in drawing conclusions, and relating conclusions to the hypothesis;

(c) ways science investigations can be managed and monitored by teachers so that formative assessment/teacher coaching are more effective, such as using checklists or predefined sets of questions;
(d) teachers' perceptions of their role in the teaching of investigations, particularly in relation to the balance between teacher guidance and developing students' independent thinking;
(e) the effect of a combination of teacher assessment and student self-assessment on the development of students' investigation competencies;
(f) explanations for students' perceptions of their learning from investigations, particularly with regard to Assertions 28-30 and 36; and
(g) the impact on students' investigation performance of teachers' beliefs about providing help for students to perform investigations (Assertion 44) and teachers' beliefs about students' abilities to reflect on their learning (Assertion 45).

Contribution of the Research

This research makes major contributions to the knowledge and understanding of students' science investigation competencies, the cognitive apprenticeship model of instruction as applied to the teaching and learning of science investigations, and the assessment of science investigations. Because a range of data sources were used the findings are credible, and as a result of the triangulation procedures they are trustworthy.

In terms of **investigation competency** numerous skills were identified as being learned including social and workplace skills such as *Working cooperatively, Attending to detail, Managing time* and *Being organised.* The research shows that students improve specific investigation competencies when they are situated in the

context of a science investigation, as distinct from the learning of skills and processes in isolation.

Difficulties that students encountered while performing investigations were documented. In particular, the way students attended to data that did not support their investigation hypothesis parallels decision making processes in other domains (Kuhn, 1992). An awareness and knowledge of potential areas of difficulties may result in improved instruction and curriculum resources. This will contribute to improved teaching and learning.

The implementation of the **cognitive apprenticeship instructional model** in a whole-class laboratory setting also contributes to our understanding of teaching and learning. Modelling whole investigations was not perceived to be an effective teaching strategy to introduce science investigations because learners often did not engage in the modelled task. Coaching, scaffolding and fading, and articulating between students contributed holistically to students' learning. Coaching and scaffolding were viewed as powerful ways to improve students' investigation competencies. Students' poor self-reflective and metacognitive skills implied that explicit instruction may be needed to develop skills in this area.

Finally, this research has contributed to our understanding of the **assessment** of investigations. The finding that students perceived that they learned more from talking with their peers about doing investigations than from their teachers, indicated that teachers did not play a major role in informal formative assessment during the investigations. For the formal formative assessment at the completion of investigations, classes experiencing student self-assessment made more modest improvements on some skills and competencies than classes that experienced teacher assessed investigations. This finding, supported by observational data, highlighted the role that formalised feedback plays in improving students' investigation competencies.

REFERENCES

Al Busaidi, R. N., Allsop, R. I. & Lock, R. J. (1992). Assessment of science practical skills in Omani 12th grade students. International Journal of Science Education, 14(3), 319-330.

Australian Education Council. (1994a). A national statement on science for Australian schools. Melbourne: Curriculum Corporation.

Australian Education Council. (1994b). Science: A curriculum profile for Australian schools. Melbourne: Curriculum Corporation.

Azmitta, M. & Perlmutter, M. (1989). Social influences on children's cognition: State of the art and future directions. Advances in Child Development and Behavior, 22, 89-144.

Baird, J. R. (1986a). Improving learning through enhanced metacognition: A classroom study. European Journal of Science Education, 8(3), 263-282.

Baird, J. R. (1986b). Learning and teaching: The need for a change. In J. R. Baird & I. J. Mitchell (Eds.), Improving the quality of teaching and learning: An Australian case study - The PEEL project (pp. 8-1[illegible] . Melbourne: Monash University Press.

Barnes, D. & Todd, F. (1977). Communication and learning in small groups. London: Routledge and Kegan Paul.

Beichner, R. J. (1990). The effect of simultaneous motion presentation and graph generation in a kinematics lab. Journal of Research in Science Teaching, 27, 803-815.

Bell, B. & Cowie, B. (1997). Formative assessment and science education: Research report of the learning in science project (assessment). Hamilton NZ: Centre for Science, Mathematics, Technology Education Research, University of Waikato.

Bennett, S. N. & Desforges, C. (1985). Recent advances in classroom research. British Journal of Educational Psychology Monograph Series, No 2. Scottish Academic Press.

Biggs, J. B. & Moore, P. J. (1993). The process of learning (3rd ed.). New York: Prentice Hall.

Black, P. J. (1990). APU science - the past and the future. School Science Review, 72(258), 13-28.

Black, P. J. (1993). Formative and summative assessment by teachers. Studies in Science Education, 21, 49-97.

Boud, D., Dunn, J. & Hegarty-Hazel, E. (1986). Teaching in laboratories. Guilford: SRHE & NFER-Nelson.

Brewer, J. & Hunter, A. (1989). Multimethod research: A synthesis of styles. Newbury Park, CA: Sage.

Broadfoot, P., James, M., McMeeking, S., Nuttall, D. & Stierer, B. (1990). Records of achievement: Report of the national evaluation of pilot schemes. In T. Horton (Ed.), Assessment debates (pp. 87-103). London: Hodder & Stoughton.

Brown, A., Campione, J., Webber, L. & McGilly, K. (1992). Interactive learning environments: A new look at assessment and instruction. In B. Gifford & M. O'Connor (Eds.), Future assessments: Changing views of aptitude, achievement and instruction (pp. 121-211). Boston: Kluwer.

Brown, J. B., Collins, A. & Duguid, P. (1989). Situated cognition and the culture of learning. Educational Researcher, 18(1), 32-42.

Brown, A. & Ferrara, R. (1985). Diagnosing zones of proximal development. In J. Wertsch (Ed.), Culture, communication and cognition: Vygotskian perspectives (pp. 273-305). Cambridge: Cambridge University Press.

Brown, J. B. & Palincsar, A. S. (1989). Guided, cooperative learning and individual knowledge acquisition. In L. B. Resnick (Ed.), Knowing, learning and instruction. Essays in honor of Robert Glaser (pp. 393-451). Hillsdale, NJ: Erlbaum.

Bruner, J. (1985). Vygotsky: A historical and conceptual perspective. In J. Wertsch (Ed.), Culture, communication and cognition: Vygotskian perspectives (pp. 21-34). Cambridge: Cambridge University Press.

Bryce, H. (1994). Openness of inquiry of laboratory formats currently used in lower secondary science. Honours thesis, Edith Cowan University, Perth, Western Australia.

Bryce, T. G. K., McCall, J., MacGregor, J., Robertson, I. J. & Weston, R. A. J. (1991). TAPS 3: How to assess open-ended practical investigations in biology, chemistry and physics. Oxford: Heinemann.

Carey, S. (1985). Conceptual change in childhood. Cambridge, MA: MIT press.

Chee, Y. S. (1995). Cognitive apprenticeship and its application to the teaching of smalltalk in a multimedia interactive learning environment. Instructional Science, 23, 133-161.

Cheung, K. C. (1994). Practical mathematics and science teaching. In T. Husen & T. N. Postlethwaite (Eds.), The international encyclopedia of education Vol 8 (2nd ed., pp. 4643-4648.). Oxford: Pergamon.

Chi, M. T. H., Feltovich, P. J. & Glaser, R. (1981). Categorisation and representation of physics problems by experts and novices. Cognitive Science, 5, 121-152.

Christensen, C. & McRobbie, C. (1994). Group interactions in science practical work. Research in Science Education, 24, 51-59.

Coles, M. & Gott, R. (1993). Teaching scientific investigations. Education in Science, September, 8-11.

Collins, A. & Brown, J. S. (1988). The computer as a tool for learning through reflection. In H. Mandl & A. Lesgold (Eds.), Learning issues for intelligent tutoring systems (pp. 1-18). New York: Springer-Verlag.

Collins, A., Brown, J. S. & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), Cognition and instruction: Issues and agendas (pp. 453-493). Hillsdale, NJ: Erlbaum.

Cook, T. D. & Reichardt, C. S. (Eds.). (1979). Qualitative and quantitative methods in evaluation research. Beverly Hills, CA: Sage.

Curriculum Council of Western Australia. (1997). Curriculum framework. Consultation draft. Perth: Curriculum Council of Western Australia.

Denzin, N. K. (1978). The research act: A theoretical introduction to sociological methods (2nd ed.). New York: McGraw-Hill.

Denzin, N. K. (1989a). Interpretive interactionism. Newbury Park, CA: Sage.

Denzin, N. K. (1989b). The research act (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Denzin, N. K. & Lincoln, Y. S. (1994). Introduction: Entering the field of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), Handbook of qualitative research (pp. 1-17). CA: Sage.

Doise, W. & Mugny, G. (1984). The social development of the intellect. New York: Pergamon.

Doran, R. I. & Tamir, P. (1992). Results of practical skills testing. Comparative Educational Review, 18, 365-392.

Driver, R., Asoko, H., Leach, J., Mortimer, E. & Scott, P. (1994). Constructing scientific knowledge in the classroom. Educational Researcher, 23(7), 5-12.

Driver, R. & Oldham, V. (1986). A constructivist approach to curriculum development in science. Studies in Science Education, 13, 105-122.

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., Johnson, P. & Gott, R. (1996). A critical point in investigative work: Defining variables. Journal of Research in Science Teaching, 33(5), 461-474.

Education Department of Western Australia. (1997). Students outcome statements: 1997 sample paper. Perth: Government of Western Australia.

Edwards, D. & Mercer, N. (1987). Common knowledge. London: Methuen.

Erickson, G. L. (1979). Children's conceptions of heat and temperature. Science Education, 63, 221-230.

Erickson, F. (1986). Qualitative research on teaching. In M. C. Wittrock (Ed.), Handbook for research on teaching (3rd ed., pp. 119-161). New York: Macmillan.

Ertmer, P. A. & Cennamo, K. S. (1995). Teaching instructional design: An apprenticeship model. Performance Improvement Quarterly, 8(4), 43-58.

Fairbrother, R. & Hackling, M. (1997). Is the answer right? International Journal of Science Education, 19(8), 887-894.

Fetterman, D. M. (1988). Qualitative approaches to evaluation in education: The silent scientific revolution. New York: Praeger.

Fielding, N. G. & Fielding, J. L. (1986). Linking data. Beverly Hills, CA: Sage.

Filstead, W. J. (Ed.). (1970). Qualitative methodology, Chicago: Markham.

Flick, U. (1992). Triangulation revisited: Strategy of validation or alternative? Journal for the Theory of Social Behaviour, 22, 175-198.

Fontana, A. & Frey, J. H. (1994). Interviewing: The art of science. In N. K. Denzin & Y. S. Lincoln (Eds.), Handbook of qualitative research (pp. 361-376). CA: Sage.

Foulds, K., Gott, R. & Feasy, R. (1992). Investigative work in science. Durham: University of Durham.

Fowler, H. W. & Fowler, F. G. (Eds.). (1964). The concise Oxford dictionary of current English (5th edn). Oxford: Oxford University Press.

Gallagher, J. J. & Tobin, K. (1987). Teacher management and student engagement in high school science. Science Education, 71(4), 535-555.

Garnett, Pamela. J., Davies, T. D. & Hackling, M. W. (1996). Working scientifically: Investigating factors affecting the pitch of a musical note. Australian Science Teachers Journal, 42(3), 50-53.

Garnett, Patrick. J., Garnett, Pamela. J. & Hackling, M. W. (1995) Refocussing the chemistry lab: A case for laboratory-based investigations. Australian Science Teachers Journal, 41(2), 26-32.

Garnett, Patrick. J. & O'Loughlin, M. O. (1989). Using practical tests to assess laboratory work in chemistry. Australian Science Teachers Journal, 35(4), 27-37.

Germann, P. J. & Aram, R. J. (1996). Student performances on the science processes of recording data, analysing data, drawing conclusions, and providing evidence. Journal of Research in Science Teaching, 33(7), 773-798.

Gipps, C. V. (1994). Beyond testing: Towards a theory of educational assessment. London: Falmer Press.

Glaser, R. (1963). Instructional technology and the measurement of learning outcomes: Some questions. American Psychologist, 18, 519-521.

Gott, R. & Duggan, S. (1995). Investigative work in the science curriculum. Buckingham, UK: Open University Press.

Guba, E. G. & Lincoln, Y. S. (1988). Do inquiry paradigms imply inquiry methodologies? In D. Fetterman (Ed.), Qualitative approaches to evaluation in education: The silent scientific revolution (pp. 89-115). New York: Praeger.

Guba, E. & Lincoln, Y. S. (1989). Fourth generation evaluation. Beverly Hills: Sage.

Hackling, M. W. (1998). Working scientifically: Implementing and assessing open investigation work in science. Resource book for primary and secondary teachers of science prepared for the Education Department of Western Australia. Perth: Edith Cowan University.

Hackling, M. W. & Fairbrother, R. W. (1996). Helping students to do open investigations in science. Australian Science Teachers Journal, 42(4), 26-33.

Hackling, M. W. & Garnett, Patrick J. (1993). Effect of context and gender on application of science investigation skills. Research in Science Education, 23, 104-109.

Hackling, M. W. & Garnett, Patrick J. (1995). The development of expertise in science investigation skills. Australian Science Teachers Journal, 41(4), 80-86

Haertel, E. (1993, April). Evolving conceptions of the generalisability of performance assessments. Paper presented at the AERA conference, Atlanta.

Hargraves, N. N. & Lynch, P. P. (1987). Studies in assessment: The practical examination revisited. Research in Science Education, 17, 202-211.

Harris, D. & Bell, C. (1994). Evaluating and assessing for learning (Rev. ed.). London: Kogan Page Ltd.

Hayes-Roth, B. & Hayes-Roth, F. (1979). A cognitive model of planning. Cognitive Science, 3, 275-310.

Hegarty-Hazel, E. (Ed.). (1990). The student laboratory and the science curriculum. London: Routledge.

Hennessy, S. (1993). Situated cognition and cognitive apprenticeship: Implication for classroom learning. Studies in Science Education, 22, 1-41.

Hodson, D. (1988). Experiments in science and science teaching. Educational Philosophy and Theory, 20(2), 53-66.

Hodson, D. (1990). A critical look at practical work in school science. School Science Review, 70(256), 33-40.

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

Hodson, D. (1993). Rethinking old ways: Towards a more critical approach to practical work in school science. Studies in Science Education, 22, 85-142.

Hodson, D. (1996). Laboratory work as scientific method: Three decades of confusion and distortion. Journal of Curriculum Studies, 28(2), 115-135.

Janesick, V. J. (1994). The dance of qualitative research design: Metaphor, methodolatry, and meaning. In N. K. Denzin & Y. S. Lincoln (Eds.), Handbook of qualitative research (pp. 209-219). CA: Sage.

Jarvela, S. (1996). Qualitative features of teacher-student interaction in a technologically rich learning environment based on a cognitive apprenticeship model. Machine-Mediated Learning, 5(2), 91-107.

Johnson, S. D. & Fischbach, R. M. (1992). Teaching problem-solving and technical mathematics through cognitive apprenticeship at the community college level. (Research report No. 143.) Berkeley, CA: National Centre for Research in Vocational Education.

Johnstone, A. H. (1981). Chemical education research - facts, findings and consequences. Chemistry in Britain, 17(3), 130-135.

Johnstone, A. H. & Letton, K. M. (1990). Investigating undergraduate laboratory work. Education in Chemistry, 27(1), 9-11.

Johnstone, A. H. & Letton, K. M. (1991). Practical measures for practical work. Education in Chemistry, 28(3), 81-83.

Johnstone, A. H. & Wham, A. J. B. (1982). The demands of practical work. Education in Chemistry, 19(3), 71-73.

Jones, A., Simon, S., Fairbrother, R., Watson, R. & Black, P. J. (1992). Development of open work in school science. Hatfield: Association for Science Education.

Karplus, R. (1977). Science teaching and the development of reasoning. Berkeley, California: University of California.

Kempa, R. F. & Ayob, A. (1995). Learning from group work in science. International Journal of Science Education, 17(6), 743-754.

Klenowski, V. (1995, April). Student self-evaluation processed: Empowering students in learner-centered contexts. Paper presented at the meeting of the American Educational Research Association, San Francisco, CA.

Kuhn, D. (1992). Thinking as argument. Harvard Educational Review, 62(2), 155-178.

Kuhn, D., Amsel, E. & O'Loughlin, M. (1988). The development of scientific thinking. New York: Hardcourt Brace Jovanovich.

Lawson, A. E. (1995). Science teaching and the development of thinking. Belmont, CA: Wadsworth.

Lincoln, Y. S. & Guba, E. G. (1985). Naturalistic inquiry. Newbury Park, CA: Sage.

Linn, M. C. & Burbules, N. C. (1993). Construction of knowledge and group learning. In K. Tobin (Ed.), The practice of constructivism in science education (pp. 91-119). Washington: AAAS Press.

Linn, R. L., Baker, E. & Dunbar, S. (1991). Complex performance-based assessment: Expectations and validation criteria. Educational Researcher, 20(8), 15-21.

Lloyd, B. W. (1992). The 20th century general chemistry laboratory. Journal of Chemical Education, 69, 866-869.

Lock, R. (1989). Assessment of practical skills. Part 1. The relationship between component skills. Research in Science and Technology Education, 7(2), 221-233.

Lock, R. (1990). Assessment of practical skills. Part 2. Context dependency and construct validity. Research in Science and Technology Education, 8(1), 35-52.

Lunetta, V. N. (1998). The school science laboratory: Historical perspectives and contexts for contemporary teaching. In B. J. Fraser & K. G. Tobin (Eds.), International handbook of science education: Vol. 2 (pp. 249-264). Dordrecht: Kluwer Academic Publishers.

Lunetta, V. N., Hofstein, A. & Giddings, G. (1981). Evaluating science laboratory skills. The Science Teacher, 48(1), 22-25.

Lynch, P. P. (1987). Laboratory work in schools and universities: Structures and strategies still largely unexplored. Australian Science Teachers Journal, 32(4), 31-39.

Mackeith, B. (Ed.). (1989). The World of Science: Vol. 17. Edinburgh: Southside.

McFee, G. (1992). Triangulation in research: two confusions. Educational Research, 34(3), 215-219.

Millar, R. (1991). A means to an end: The role of processes in science education. In B. E. Woolnough (Ed.), Practical Science (pp. 43-52). Milton Keynes: Open University Press.

Mines, G. (1995). Sci 1 investigations: Management and assessment at key stage 4. Education in Science, 165, 13-15.

Murphy, P. (1988). Insights into pupils' responses to practical investigations from the APU. Physics Education, 23(6), 330-336.

National Academy of Sciences & National Research Council. (1996). National science education standards. Washington, DC: National Academy Press.

Novak, J. (1988). Learning science and the science of learning. Studies in Science Education, 15, 77-101.

Nussbaum, J. & Novick, S. (1982). Alternative frameworks, conceptual conflict and accommodation: Towards a principled teaching strategy. Instructional Science, 11, 183-200.

Palinscar, A. S. & Brown, A. L. (1984). Reciprocal teaching of comprehension - fostering and monitoring activities. Cognition and Instruction, 1, 117-175.

Patton, M. Q. (1981). Practical evaluation. CA: Sage.

Patton, M. Q. (1986) Utilization-focused evaluation (2nd ed.). Beverly Hills, CA: Sage.

Patton, M. Q. (1988). Paradigms and pragmatism. In D. M. Fetterman (Ed.), Qualitative approaches to evaluation in education: The silent scientific revolution (pp. 116-137). New York: Praeger.

Patton, M. Q. (1990). Qualitative evaluation and research methods (2nd ed.). London: Sage.

Piaget, J. (1972). Psychology and epistemology: Towards a theory of knowledge. London: Allen Lane the Penguin Press.

Pieters, J. M. & DeBruijn, H. F. M. (1992). Learning environments for cognitive apprenticeship: From experience to expertise. In P. A. M. Kommers, D. H. Jonassen & T. Mayes (Eds.), Cognitive tools for learning (pp. 241-248). Berlin: Springer-Verlag.

Polya, G. (1945). How to solve it: A new aspect of mathematical methods. Princeton, NJ: Princeton University Press.

Popham, J. (1984). Specifying the domain of content or behaviours. In R. A. Berk (Ed.), A guide to criterion referenced test construction (pp. 29-48). Baltimore, MD: John Hopkins University Press.

Popham, J. (1992, April). Educational testing in America: What's right, what's wrong? Paper presented at the AERA conference, San Francisco.

Radnor, H. & Shaw, K. (1995). Developing a collaborative approach to moderation. In H. Torrance (Ed.), Evaluating authentic assessment (pp. 124-143). Buckingham, UK: Open University Press.

Reif, F. & Larkin, J. H. (1991). Cognition in scientific and everyday domains: Comparison and implications. Journal of Research in Science Teaching, 28, 733-760.

Renner, J. (1982). The power of purpose. Science Education, 66(5), 709-716.

Richmond, G. & Striley, J. (1996). Making meaning in classrooms: Social processes in small-group discourse and scientific knowledge building. Journal of Research in Science Teaching, 33(8), 839-858.

Rigano, D. L. & Ritchie, S. M. (1995). Student disclosures and fraudulent practices in school laboratories. Research in Science Education, 25(4), 353-363.

Roberts, D. A. (1996). What counts as quality in qualitative research? Science Education, 80(3), 243-248.

Ross, J. A. & Robinson, F. G. (1987). The use of rule structures in teaching experimental design to secondary school students. Science Education, 71(4), 571-589.

Roth, W. M. (1995) Authentic school science: Knowing and learning in open-inquiry science laboratories. Netherlands: Kluwer Academic Publishers.

Roth, W. M. & Bowen, G. M. (1995). Knowing and interacting: A study of culture, practices, and resources in grade 8 open-inquiry science classroom guided by a cognitive apprenticeship metaphor. Cognition and Instruction, 13(1), 73-128.

Roth, W. M. & McGinn, M. K. (1997). Graphing: cognitive ability or practice? Science Education, 81(1), 91-106.

Roth, W. M. & Roychoudhury, A. (1993). The development of science process skills in authentic contexts. Journal of Research in Science Teaching, 30(2), 127-152.

Rowell, J. A. & Dawson, C. J. (1984). Controlling variables: Testing a programme for teaching a general solution. Research in Science and Technological Education, 2(1), 37-46.

Roychoudhury, A. & Roth W. M. (1996). Interactions in an open-inquiry physics laboratory. International Journal of Science Education, 18(4), 423-445.

Sadler, R. (1987). Specifying and promulgating achievement standards. Oxford Review of Education, 13, 191-209.

Scardamalia, M. & Bereiter, C. (1985). Fostering the development of self-regulation in children's knowledge processing. In S. F. Chipman, J. W. Segal & R. Glaser (Eds.), Thinking and learning skills: Research and open questions (pp. 563-577). Hillsdale, NJ: Lawrence Erlbaum Associates.

Schoenfeld, A. H. (1985). Mathematical problem solving. Orlando, FL: Academic Press.

Sharp, R. & Green, A. (1975). Education and social control: A study in progressive primary education. London: Routledge and Kegan Paul.

Shavelson, R. J. & Baxter, G. P. (1992). What we've learned about assessing hands-on science. Educational Leadership, 49(8), 20-25.

Siegel, H. (1988). Educating reason: Rationality, critical thinking and education. New York: Routledge.

Solomon, J. (1988). Learning through experiment. Studies in Science Education, 15, 103-108.

Tamir, P. & Doran, R. L. (1992). Conclusions and discussions of findings related to practical skills testing in science. Comparative Educational Review, 18, 393-406.

Tikunoff, W. & Ward, B. (1980). Interactive research and development on teaching. San Francisco: Far West Laboratory for Educational Research and Development.

Tobin, K. (1990). Research on science laboratory activities: In pursuit of better questions and answers to improve learning. School Science and Mathematics, 90(5), 403-418.

Tobin, K. G. & Gallagher, J. J. (1987). What happens in high school science classrooms? Journal of Curriculum Studies, 19, 549-560.

Toh, K. A. (1991). Factors affecting success in science investigations. In B. Woolnough (Ed.), Practical science (pp. 89-100). Milton Keynes: Open University Press.

Toh, K. A. & Woolnough, B. E. (1990). Assessing, through reporting, the outcomes of scientific investigations. Educational Research, 32(1), 59-65.

Torrance, H. (1993). Formative assessment: Some theoretical problems and empirical questions. Cambridge Journal of Education, 23(3), 333-343.

Torrance, H. (1995). Introduction. In H. Torrance (Ed.), Evaluating authentic assessment (pp. 1-8). Buckingham, UK: Open University Press.

United Kingdom Department of Education and Science. (1988). Science 5-16: A statement of policy. London: HMSO.

Volet, S. E. (1991). Modeling and coaching of relevant metacognitive strategies for enhancing university students' learning. Learning and Instruction, 1, 319-336.

von Glaserfeld, E. (1989). Cognition, construction of knowledge and teaching. Synthese, 80, 121-140.

Vygotsky, L. (1986). Thought and language. Cambridge, MA MIT press.

Wavering, M. J. (1989). The logical reasoning necessary to make line graphs. Journal of Research in Science Teaching, 26, 373-379.

Wood, R. (1986). The agenda for educational measurement. In D. L. Nuttal (Ed.), Assessing educational achievement (pp. 185-204). London: Falmer Press

Woolnough, B. E. (1989). Towards a holistic view of processes in science education. In J. J. Wellington (Ed.), Skills and processes in science education: A critical analysis (pp. 115-134). London: Routledge.

Woolnough, B. E. (1991). Practical science as a holistic activity. In B. E. Woolnough (Ed.), Practical science (pp. 181-187). Milton Keynes: Open University Press.

Woolnough, B. & Allsop, T. (1985). Practical work in science: Cambridge Science Education Series. Cambridge: Cambridge University Press.

Woolnough, B. E. & Toh, K. A. (1990). Alternative approaches to assessment of practical work in science. School Science Review, 71(256), 127-131.

APPENDIX A

STUDENT OUTCOME STATEMENTS: STANDARDS FOR THE ASSESSMENT OF THE PRETEST AND POSTTEST INVESTIGATIONS

Level	Planning investigations	Conducting investigations	Processing data	Evaluating the investigation
FOS	The student demonstrates an awareness of, explores and responds to objects and events in their environment using their senses. Is able to make choices based on their experience			
1	The student focuses on a problem, responds to teacher's suggestions and carries out simple activities requiring observation and sharing of observations			
SOS	Focuses on problem and responds to teachers suggestions and questions	Carries out activities involving a small number of steps. Observes and describes	Students share observations	Not applicable
Pointers	Tells about their own experiences of a phenomenon Contributes to a list of what the class members know about a topic Responds to teacher questions about "What would we like to know?", "How can we find out?"	Collects materials following teacher directions Carries out an investigation involving a small number of steps. Tells what they have done and observed	Tells what they observed Acts out what they did or what happened Draws what happened.	Not applicable
2	The student, when given a focus question in a familiar context contributes elementary ideas about variables and procedures, collects and makes limited records of data and can say whether what happened was expected.			
SOS	Given a focus question in a familiar context can identify some of the variables to be considered	Observes, classifies, describes, makes simple non-standard measurements and limited records of data. Independent variables are usually discrete.	Makes comparisons between objects or events observed.	Comments on what happened and can say whether what happened was expected
Pointers	Contributes to brainstorming of variables that could be considered. Recognises which variable is to be the focus of the investigation. Says how information will be gathered about the variable.	Uses simple non standard measurements Uses pictures or words or numbers to record observations. Places objects into groups.	Compares events eg the marble rolled further on the steeper slope Compares places eg there were more animals in the bush than in the park.	Comments on what happened Can say whether what happened was expected. Says that the outcome was different from the prediction.
3	The student shows some awareness of the need for fair testing and makes simple predictions. Collects and organises numerical data and descriptive information using simple tables, graphs, diagrams and lists of observations. Identifies main features and patterns in the data and identifies difficulties experienced in conducting the investigation.			
SOS	Planning for investigations show some awareness of the need for fair testing. Makes simple predictions (not guesses) based on personal experience.	Simple equipment is used in a consistent manner. Records data in simple tables, diagrams or as lists of observations.	Displays numerical data as tables or bar graphs, identifies patterns in data and summarises the data.	Identifies difficulties experienced in doing the investigation.
Pointers	Says that something that will be kept the same or done in the same way. From past experiences, says what they think will happen.	Uses the equipment in the same way for different trials or treatments. Makes simple measurements using standard units. Chooses forms of data presentation that are appropriate for the types of data eg lists, tables, diagrams, audio or videotape. Takes some responsibility to ensure safety.	Organises numerical data into tables. Draws simple bar graphs Identifies patterns (groups, trends or relationships) in the data. Conclusions summarise but do not explain the patterns in the data. Can relate an effect to a cause	Says that it was difficult to make exact measurements. Describes external factors that influenced the results Describes difficulties experienced in using the equipment.
4	The student plans and conducts different types of investigations taking account of the main variables. Collects data using repeat trials or replicates, identifies, summarises and explains patterns in data and makes general suggestions for improving the investigation.			
SOS	Identifies the variable to be changed, the variable to be measured and at least one variable to be controlled or, in a descriptive study, can plan for the types of observations that need to be made.	Care is taken with data collection so that data are accurate. Uses repeated trials or replicates. Independent variables are usually continuous.	Calculates averages from repeated trials or replicates. Data plotted as line graphs where appropriate Conclusions summarise and explain patterns in the data.	Makes general suggestions for improving the investigation
Pointers	Written plans specify the two main variables (independent and dependent). Written plans name at least one variable that will be kept the same. Says how data will be collected for the independent and dependent variables. Selects an appropriate data collection procedure.	Uses equipment correctly and consistently. Follows directions to correctly perform simple chemical tests. Measurements are accurate to one scale division. Makes more than one measurement for each treatment. Recognises the need for safety precautions eg wears safety glasses.	Calculates averages from repeat trials or replicates. Sums data over intervals eg daily rainfall over a month. Plots bar graphs and simple line graphs. Summarises the data and attempts to explain the patterns in the data and/or relationships between the variables.	Says that better equipment was needed to do the experiment properly. Says that measurements need to be more exact. Says that the testing needs to be done more times or more observation should have been made.
5	The student analyses a problem and formulates a plausible relationship to investigate using experimental techniques including the control of several variables and the use of preliminary trials to improve the procedure or measurement techniques. Develops scientific explanations that are consistent with the data and makes specific suggestions for improving the investigation.			
SOS	Analyses problem, formulates a question or hypothesis for testing, and plans an experiment in which several variables are controlled.	Chooses equipment that is appropriate for the task. Preliminary trials of the experimental procedure are used to improve the procedure or measurement techniques.	Conclusions are consistent with the data and explain patterns in the data in terms of scientific knowledge.	Suggests specific changes that would improve the techniques used or the design of the investigation.
Pointers	Writes a question or hypothesis to focus the planning of their investigation. Lists variables possibly important in the investigation and plans to control several of these. Plans data collection procedures, equipment and techniques to be used.	Selects an appropriate size measuring cylinder or spring balance that will enhance accuracy. Uses preliminary trials to improve the procedure or measurement technique. Takes enough measurements to gauge reliability.	Explains the patterns in the data or relationships between the variables in terms of scientific knowledge. Writes conclusions that reflect closely the magnitudes and patterns in the data.	Says how the measurement procedure can be made more accurate. Identifies a variable that was not kept the same across treatments and says how it should have been done. Says how the measurement procedure could have been applied more consistently.

Level	Planning investigations	Conducting investigations	Processing data	Evaluating the investigation
6	The student uses scientific knowledge to analyse a problem, identify variables and formulate questions for investigation. Develops methods that provide accurate and consistent information which can be used to evaluate the questions. Recognises inconsistencies in the data and suggests ways of reducing error.			
SOS	Analyses problem, formulates question or hypothesis for testing, uses scientific knowledge to identify main variables to be considered and make predictions, and plans for accurate measurement	Decides what is needed and requests equipment for the investigation Selects apparatus and instruments that enhance safety and accuracy of measurements and observations	Selects type of graph and scales that display data effectively. Conclusions are consistent with the data, explained in terms of scientific knowledge and related to the question, hypothesis or prediction	Recognises inconsistencies in the data, identifies the main sources of error and can suggest improvements that would reduce the source of error
Pointers	Uses scientific knowledge to identify the key variables that influence the phenomenon Uses scientific knowledge in developing predictions Uses scientific knowledge to select an appropriate chemical test Considers how to enhance the accuracy of measurements	Chooses equipment that enhances safety and accuracy of measurement Recognises difficulties in making accurate measurements. Uses operational definitions to enhance consistency of measurement decisions Avoids parallax error	Makes appropriate decisions about the type of graph to use which is best for the purpose and type of data Selects appropriate origin, range and intervals for graph scales Evaluates the question or hypothesis in terms of the data	Recognises that differences in observations or measurements for repeat trials or replicates are too large and represent error Identifies the main source of error Suggests changes to the design or technique that would minimise or eliminate that error
7	The student identifies a significant issue for investigation by researching the literature and develops systematic procedures for collecting accurate and consistent data. Draws credible conclusions that are consistent with their own and other data, recognises limitations, acknowledges sources of error and from this proposes improvements that would reduce these errors			
SOS	Students identify their own real-world problem for investigation, use reference material in developing an understanding of the problem, and plans one or more experiments in an ongoing investigation.	Makes systematic observations and measurements with precision using standardised techniques and recognises when to repeat measurements	Conclusions are consistent with the data, explain them in terms of scientific knowledge and are not overgeneralised. Students start to question whether the data are sufficient to support the conclusions drawn	Recognises sources of measurement error, limitations in sampling and inadequacies in control of variables and can explain how these deficiencies can be remedied
Pointers	Identifies a real world problem worthy of scientific investigation. Uses reference material and own scientific knowledge to develop an understanding of the problem. Formulates several questions or hypotheses for testing. Plans a sequence of experiments or a long term investigation. Plans a sequence of chemical tests for analytical work.	Develops or refines measurement and/or observational techniques. Consistency in measurement procedure produces precise measurements Monitors consistency of data as it is collected Makes objective decisions for discarding discrepant results and repeating measurements.	Writes conclusions that do not go beyond the data ie are not overgeneralised Writes conclusions which indicate an appropriate level of confidence in the data.	Writes about limitations in measurement, control of variables and sampling in the conclusions section of his/her laboratory report Reviews limitations and main sources of error when presenting a seminar on his/her investigation
8	The student's planning shows an *a priori* recognition of the need for control of variables, accuracy of measurement, adequate sample size and repeated trials or replication. Students critically analyse procedures and data to produce accurate, valid and reliable information to guide the on-going refinement of investigation designs and techniques and generate credible explanations for phenomena being investigated			
SOS	Working on their own problem students planning shows an *a priori* recognition for the need for control of variables, accuracy of measurement, adequate sample size and repeated trials or replication.	Makes judgements about the accuracy required, range and intervals of measurement, and decides what observations are necessary and sufficient in qualitative work.	Identifies anomalous observations and measurements and allow for these when graphs and conclusions are drawn. Conclusions show awareness of uncertainty in data, are not overgeneralised and include a discussion of limitations of the methods of data collection and/or design.	Evaluates the findings and the experimental design, reformulates the problem, plans follow-up experiments in an on-going investigation, and refinements to experimental techniques and design
Pointers	Understanding of the need for thorough control of variables influences the search for important variables and ways of controlling them in the design and techniques to be used. Planning for measurement considers the range and intervals of measurement, choice of apparatus and techniques, and refinement of the measurement technique. Plans for data collection show a need for samples to be representative and sufficiently large, and for repeat trials or replication to be used where appropriate. Plans for triangulation of data.	Chooses appropriate range and intervals of measurement for the independent variable. Different measurements are made to an appropriate level of accuracy Different measurements are recorded with an appropriate number of significant figures. Takes enough readings to be able to estimate the error of measurement. Collects data in ways that allow for triangulation and checks on consistency. Minimises impact of the investigation on animals, others and the environment.	Accounts for anomalous observations when graphing and/or interpreting data Writes conclusions that include a discussion of the limitations of the design and/or data collection methods.	Evaluates the impact of the investigation on animal subjects and the environment Evaluates the findings in terms of the hypothesis, existing theory and the problem Reformulates the problem Evaluates the design and procedure and identifies changes that are needed for follow-up experiments

APPENDIX B

TEACHING ROSTER: CLASSES AND ACTIVITIES

Starting Date	Class TN	Class TC	Class SC
May	**Mrs Cross** TOSIS Worksheet on terminology Pretest Lipase investigation Teacher modelled acid and carbonate investigation Catalyst investigation	**Miss Mills** TOSIS Worksheet on terminology Pretest Lipase investigation Teacher modelled acid and carbonate investigation Catalyst investigation	**Mrs Grant** TOSIS Worksheet on terminology Pretest Lipase investigation Teacher modelled pitch of a closed pipe investigation Ukulele investigation
August	**Mrs Grant** Teacher modelled pitch of a closed pipe investigation Ukulele investigation Worksheet on marking an investigation Electromagnet investigation	**Mr Brogo** Teacher modelled pitch of a closed pipe investigation Ukulele investigation Worksheet on marking an investigation Electromagnet investigation	**Mrs Cross** Teacher modelled acid and carbonate investigation Catalyst investigation Worksheet on marking an investigation Panadol investigation
November	**Mrs Cross** Panadol investigation Posttest Trypsin investigation	**Miss Mills** Panadol investigation Posttest Trypsin investigation	**Mrs Grant** Electromagnet investigation Posttest Trypsin investigation

APPENDIX C

TYPICAL LESSON SEQUENCE

Dates and lesson sequence for Class SC

	Lesson sequence
Lesson 1 Tue, May 2nd	**Pretest** Test of Science Investigation Skills (TOSIS)
Lesson 2 Wed, May 3rd	**Worksheet 1:** Terminology of investigations
Lesson 3 Thurs, May 4th	**Pretest Investigation: Lipase investigation** Planning and trialing
Lesson 4 Fri, May 5th	Lipase investigation continued Conducting the investigation (data gathering)
Lesson 5 Mon, May 8th	Lipase investigation continued Processing data and writing up the investigation
Lesson 6 Tue, May 9th	**Assessment feedback:** Lipase investigation **Teacher modelled investigation:** Pitch of a closed pipe
Lesson 7 Wed, May 10th	Teacher modelled investigation continued
Lesson 8 Thurs, May 11th	**Ukulele investigation** Planning and trialing
Lesson 9 Fri, May 12th	Ukulele investigation continued Conducting the investigation (data gathering)
Lesson 10 Mon, May 15th	Ukulele investigation continued Processing data and writing up the investigation
Lesson 11 Tue, May 16th	**Assessment feedback:** Ukulele investigation

APPENDIX D

INVESTIGATION PLANNING AND REPORT SHEETS

(CONDENSED VERSIONS)

Pretest Investigation: Lipase Investigation

Date: ______________________ Name: __________________________

Other members in the group: ____

Problem

Lipase is an enzyme found in the small intestine. Lipase helps to digest food by breaking down fats into fatty **acids.** For example it breaks down the fat in milk into fatty **acids.** The environment in the small intestine, where the lipase works, is very slightly **basic** and the temperature is **37 ^{o}C**. The fatty acids that are produced by the action of lipase make the solution acidic.

What variables might affect how quickly this will take place?
Which variables will you investigate?
Write an hypothesis for your investigation.

Planning

Describe your plan for the investigation.

Preliminary Trials

Describe any trials you did before starting to collect your data.
Describe any changes that you made to your plan as a result of your preliminary trials.

Collecting Data

Describe how you made sure that your data were accurate.
Record your data in a table.

Processing Data

If appropriate draw a graph of your data.
What conclusion(s) can you make from your results?
Explain why you can make that conclusion.

Evaluating the Findings

What are the main sources of error in your experiment?
How confident are you about the conclusion(s)?
If you were to do this again how would you change the investigation?

Background Information for the Pretest Investigation: Lipase Investigation

Safety: Wear safety glasses and disposable gloves. Take care with hot water.

Procedure

Measure 2.5 mL of full cream milk into a test-tube (labelled A)

Prepare your other samples of milk. What you use depends on **your** investigation.

2% lipase solution is provided. Measure 4 mL of this solution into test tube A. What is added to the other tubes will depend on your investigation but you should add 4 mL of liquid to each test tube.

Add a small amount of Universal indicator solution to each test tube, shake gently to mix and observe the colour in each test tube. 1 mL is approximately 20 drops.

Unless you are investigating how temperature affects the reaction place the test tubes in warm water (at approximately 37 °C), during the investigation. Make your observations every 2 minutes.

Equipment

for each group placed in nine trays which can be kept for several lessons

- 6 test tubes and test tube rack
- 150 mL of 2% lipase solution
- Universal indicator
- glass rod
- thermometer
- 10 mL syringe to measure milk
- 10 mL syringe to measure lipase
- full cream milk. Put your name on the milk sample because it will need to be returned to the fridge.
- sticky labels
- 2 x 100 mL beakers, one for lipase and one for milk
- rubber gloves
- distilled water
- stop clock
- 4 hot water baths set at 37 °C. Do not change the temp setting of the water baths once the experiment has commenced.
- marker for writing on test tubes
- safety glasses
- graph paper

Depending on what you decide to investigate you may need the following

for groups investigating the <u>amount of fat in the milk</u>

- different milk samples such as buttermilk, full cream milk, Hilo milk and powdered skim milk that is made up. (Do not use fresh skim milk as it may have non-milk fat added and will produce misleading results.)

for groups measuring the <u>temperature of the reaction</u>

- large beakers for water baths
- ice
- hot water from the hot water tap
- very hot water from the urn (85 °C)

Posttest Investigation: Trypsin Investigation

Date: ______________________ Name: ___________________________

Other members in the group: ____

The Problem

Trypsin is an enzyme found in the small intestine. Trypsin helps to digest food by breaking down protein. The environment in the small intestine, where the trypsin works, is very slightly **basic** and the temperature is **37 ^{o}C**. Trypsin, for example, breaks down the protein called gelatin. Gelatin is found on undeveloped photographic film and when a film is put in a trypsin solution the gelatin reacts.

What variables might affect how quickly this reaction will take place?
Which variables will you investigate?
Write an hypothesis for your investigation.

Planning

Describe your plan for the investigation.

Preliminary Trials

Describe any trials you did before starting to collect your data.
Describe any changes that you made to your plan as a result of your preliminary trials.

Collecting Data

Describe how you made sure that your data were accurate.
Record your data in a table.

Processing Data

If appropriate draw a graph of your data.
What conclusion(s) can you make from your results?
Explain why you can make that conclusion.

Evaluating the Findings

What are the main sources of error in your experiment?
How confident are you about the conclusion(s)?
If you were to do this again how would you change the investigation?

Background Information for the Posttest Investigation: Trypsin Investigation

Safety: Wear safety glasses and disposable gloves. Take care with hot water.

Background information

Trypsin is an enzyme that reacts with protein. A photographic film is used as the source of protein. Photographic films are coated with a protein called gelatin. When the protein is broken down by trypsin the film becomes clear. Remember that the temperature of the human body is 37 °C.

You will only be allowed **12 pieces** of photographic film. Use forceps to hook these onto plastic covered wires.

As a starting point use a syringe to measure 4 mL of trypsin solution and into a test tube. Then use another syringe to measure 1 mL of sodium hydroxide solution into the test tube.

The reaction will start when you add the film to the liquid. Immediately place the test tube into a water bath at 37 °C. Make observations every 2 minutes.

What you do now depends on what your group has decided to investigate.

It is important to remember that the trypsin is extremely expensive and because of this you will only be allowed a total of **50 mL**. This means that you should **never use more than 5 mL** in any one trial. Most groups will use **4 mL of trypsin and 1 mL of sodium hydroxide** each time.

Equipment

for each group placed in 9 trays which can be kept for several lessons

- 6 test tubes in a test tube rack
- 50 mL 2 % trypsin in a 100 mL beaker, labelled
- 50 mL 0.1 M sodium hydroxide solution in a 100 mL labelled beaker,
- 2 syringes
- graph paper
- 6 wires and 12 pieces of photographic film
- permanent marker pen
- access to water bath at 37 °C
- thermometer
- stop clock
- forceps

for groups investigating the amount of trypsin
- as above

for groups investigating temperature

- 4 large beakers for water baths
- ice
- water from the hot water tap
- water from the urn at approximately 85 °C

for groups investigating the solvent in which trypsin works

- 0.01 M and 0.2 M NaOH concentration
- 0.1 M HCl solution
- NaCl solution

APPENDIX E
THE INVESTIGATIONS

The Panadol Investigation

Medicines need to be dissolved so that they can be absorbed into the blood stream. Tablets that dissolve quickly have the potential to be absorbed more rapidly and will start to have an effect sooner than those that take a long time to dissolve. You will investigate **what might affect the rate at which Panadol Clear dissolves?**

The Catalyst Investigation

Some contact lenses are sterilised with hydrogen peroxide. A metal disc acts as a catalyst to speed up the process. The metal disc catalyst decomposes the hydrogen peroxide into water and oxygen gas. When the contact lens wearer inserts the clean lens into their eye there is no hydrogen peroxide left. Some metal compounds can also be used to speed up the rate at which hydrogen peroxide decomposes. You will investigate one factor **which might affect the rate at which the hydrogen peroxide reacts to form oxygen gas and water?**

The Ukulele Investigation

All musical instruments need to be tuned. Stringed instruments such as guitars and violins always need tuning before they are played. This ensures that the pitch, which is the highness or lowness of a note is correct. You will **investigate what might affect the pitch of a stringed instrument.**

The Electromagnet Investigation

Electromagnets are used to separate scrap iron and other magnetic metals from rubbish yards and tips. You will **investigate what might affect the strength of an electromagnet.**

APPENDIX F

SCAFFOLDED INVESTIGATION PLANNING AND REPORT SHEET (CONDENSED VERSION)

The Problem

Statement of the problem to be investigated (see Appendix D)

In this space write down anything that you already know about this topic.
List the variables that could be investigated. A variable can be changed or kept constant. Variables need to be measured.

Planning
Decide what **you** will investigate and write an hypothesis. An hypothesis should state the relationship between the independent and dependent variables.
The independent variable is ...
The dependent variable is ...
Complete the table for this investigation.

	Variable/s	Unit/s	How the variable will be measured
Independent			
Dependent			
Factors kept constant (controlled variables)			

Briefly outline what you plan to do to test the hypothesis. Remember that you may need to modify your plan after your preliminary trials. You will be required to write a full description of your method later.

Preliminary trials
Conduct some preliminary trials to try out your plan.
Describe what you learned from these initial trials **and** describe modifications you made to your initial plan.

Collecting data
In detail, describe how you conducted the experiment. This description should be sufficiently detailed for someone else to follow and obtain similar results. It should include a diagram drawn in pencil.
Record the data in a table

NB *Most novice investigators take an insufficient number of readings when compared with experts.*

Independent variable (units)	Dependent variable (units)			
	Trial 1	Trial 2	Trial 3	Average

TITLE: ______________________

Processing data
Draw a graph on the graph paper or a chart to find if there is a pattern in the results.
Note: Plot the independent variable on the horizontal axis.
What is the relationship between the variables that you have investigate?
What conclusion(s) can you make from the results? Was the hypothesis supported?

Evaluating the findings
What were the main sources of error? (sample size and selection, measurement error, control of variables)
How confident are you about the conclusion(s)?
If you were to do this again how would you improve the design of your investigation?
What have you learned from your investigation about:

the problem/phenomenon
the way science investigations are carried out
how ideas in science develop

What parts of the investigation did you find easy and what parts were hard?
Did your group work well together? If not then say how the group could work better.
Have you any other comments to make about the investigation.

Paste your graph here

APPENDIX G

THE WORKSHEETS (CONDENSED VERSIONS)

Worksheet on Terminology

In this worksheet you will revise some of the ideas about investigations that you studied in Year 8. Do you remember the terms hypothesis, variable, independent variable and dependent variable?

Worked example
Sarah is making an apple pie. The raw apples she puts into the pie are quite firm. She wonders "What happens to apples when they are cooked for a long time?"

Hypothesis:	The longer the apple is heated **Independent variable** heating time	the	softer it will become. **Dependent variable** softness of apple

Notice how the hypothesis refers to both the independent and dependent variables.

Independent variable: This is the **time** that the apple is heated. This can be measured with a clock. Remember that this is the variable that you deliberately change.

Dependent variable: This is the **softness of the apple**. This is the variable that changes in response to the independent variable.
Can you suggest a way to judge the softness of the apple?

Controlled variables: These must be kept constant because they might affect the dependent variable. They include the **method of heating, temperature, type of apple, size of the apples.**
How will you heat the apple? In boiling in water, in an electric oven, in a microwave oven etc. This needs to be controlled for all experiments. Why?
To what temperature do you heat the apple? This needs to be controlled? Why?
The apples need to be of the same type. Why?
The apple pieces need to be the same size. Why?

Example 1: The bread investigation

Kate wanted to see if bread tasted better with more salt in it. She decided to bake four loaves of bread with different amounts of salt in them. Kate baked the bread in the same oven, for the same time in the same baking tins. She also used the same amount of the other ingredients in each loaf. Kate recorded her results in the following way:

Table 1: The effect of salt content on the taste of bread

LOAF	SALT (g)	TASTE
1	5	bad
2	10	all right
3	15	good
4	20	great

What is Kate's hypothesis?
What are the variables in this experiment?

'The independent variable is the variable that you change. It changes in the dependent variable.
What is the independent variable in Kate's experiment?
'The dependent variable changes in response to the independent variable.
What is the dependent variable in the bread experiment?
Controlled variables are the variables that are kept constant to make the test fair.
What variables does Kate keep the same (controlled) in her experiment?

Example 2: Alcohol and urine output

Alcohol affects the body in many ways. One of the effects is that it makes people pass more water as urine than they normally would. An investigator set up an experiment to test the effect of alcohol on the amount of urine passed out of the body. A large number of people were involved in the experiment. Each person drank enough alcohol to reach the blood alcohol readings listed. Their rate of urine loss was then measured. The experiment was performed a number of times with each person. An average was then calculated.

Table 2: Effect of blood alcohol on the rate of passing urine from the body

Alcohol reading in blood (parts alcohol per million parts of blood)	Rate of urine passed out of the body (millilitres of urine each hour)
0	20
0.2	35
0.5	50
0.8	100

Example 3: The reaction hydrochloric acid with a metal

Acids can be harmful because they react quickly and are corrosive. With many metals hydrochloric acid reacts to release bubbles of hydrogen gas. A Year 9 student set up an experiment to see how the concentration of hydrochloric acid affects its rate of reaction with magnesium metal. A large number of tests were conducted for each different concentration of acid. In each test 1.5 grams of magnesium metal was used. The time taken for all the magnesium to react was recorded. Tests were carried out a number of times at each different acid concentration. An average was then calculate.

Table 3: Effect of hydrochloric acid concentration on reaction rate with magnesium

Concentration of acid (moles of acid per litre)	Rate of reaction of magnesium metal (minutes taken for all the metal to react)
2.0	4
3.0	3
4.0	2
5.0	1

In this experiment:
The hypothesis being tested is
The independent variable is:
The dependent variable is:
Which variables should have been controlled?

Example 4: Drying clothes

A student noticed that when she went to collect her washing off the clothesline, her black tee shirt was always dry and her light coloured clothes were often damp. This observation intrigued her and she decided to investigate.
In her experiment:
An hypothesis she could test is:
The independent variable would be:
The dependent variable would be:
Which variables should have been controlled?

Example 5: The effects of moisturiser on skin

Cross-linked elastin is an expensive moisturiser and a cosmetic company claims that the more of it you use the fewer wrinkles you get. You doubt the claims of the cosmetic company and decide to do your own investigation to test its "magical" properties.
In your experiment:
The hypothesis being tested is:
The independent variable is:
The dependent variable is:
Which variables should have been controlled?

Worksheet on Marking an Investigation (Condensed version)

Your teacher is ill and away from school. The marks for the reports are due in and you have been asked to help out by marking some science investigations. **Your task is to work out a marking plan for the investigations.** To do this you first need to work out what things should be included in the students' investigations and then how you will allocate the marks. Using the headings Planning, Preliminary Trials, Collecting Data, Processing Data and Evaluating the Findings, write down the things that you will be looking for, and your marks allocation.

You will notice that there is another column, called "Actual Mark". This column is for you to use when you mark the investigation that follows.

PLANNING: Things that may be included	Marks allocation	Actual Mark
PRELIMINARY TRIALS: Things that may be included		
COLLECTING DATA: Things that may be included		
PROCESSING DATA: Things that may be included		
EVALUATING THE FINDINGS: Things that may be included		
TOTAL MARK:		

Note: Four graphs are omitted from the students' report that follows. They drew bar graphs for each set of data to obtain pre and post exercise graphs for both the asthmatics and non asthmatics. Each subject was plotted as a bar.

Use the marking plan and to mark the investigation and write comments on the performance of the students.

The following investigation was performed by a group of four students. The students' investigation has been typed although they presented the investigation as a poster. **Your task is to mark the investigation using the marking sheet that you have designed and to write some comments**. Enter the marks in the column called "Actual Marks" and write your comments in the box allocated.

Investigation into Heart Beat — *by Jasmine, Nicole, Katherine & Sarah*

The problem *My group consisting of Jasmine, Nicole and Katherine and myself are investigating the different pulse rates between asthmatics and non-asthmatics. We are investigating if exercise has any affect on the pulse rate. We will do this by measuring the pulse rate for two exercises.*

The hypothesis *Asthmatics will have a higher pulse rate after exercise than non- asthmatics.*

The independent variable: *Whether students were asthmatic or non-asthmatics.*

The dependent variable *The pulse rate.*

Factors kept constant *The clock with which we measured the time, the time of the day, all results were done in Period 6. The same exercises, the same thermometer. We went in pairs so the same person recorded the same subject.*

Use of equipment *stairs, clock, thermometer and the floor*

Pulse rates between asthmatics and non asthmatics

		Ex 1 B	Ex 1 A	Ex 1 Change	Ex 2 B	Ex 2 A	Ex 2 Change
Melissa	A	80	120	40	80	140	60
Katie	A	84	152	68	92	113	24
Davina	A	60	118	58	90	146	56

Average change = 55.3 (Ex 1) — *Average change = 46.6* (Ex 2)

		Ex 1 B	Ex 1 A	Ex 1 Change	Ex 2 B	Ex 2 A	Ex 2 Change
Jasmine	NA	82	148	66	88	108	20
Meg	NA	80	114	34	90	112	22
Angie	NA	74	138	64			
Karlia	NA	80	110	30	70	120	50
Jane	NA	56	130	74	92	114	52
Libby	NA	90	94	4	78	62	-16
Claire	NA	80	95	15	70	112	42
Judy	NA	94	94	0	68	98	30
Trisha	NA	64	100	36			
Harriot	NA	94	130	36	86	130	44
Ricki	NA	60	70	10	48	62	140

Average change = 33.54 (Ex 1) — *Average change = 28.6* (Ex 2)

Difficulties with the investigation

The difficulties with our investigation were people being absent and miscalculations.
Our results show that asthmatic students have a much higher pulse rate than non asthmatics.
The average change in pulse for an asthmatic was 55.3 and for the non-asthmatics 33.5

We learnt from our investigation that our hypothesis was true.

APPENDIX H

TEACHER MODELLED INVESTIGATIONS

Teacher Modelled Acid and Carbonate Investigation

Marble statues and limestone buildings consist of the chemical compound calcium carbonate. Acid solutions react with calcium carbonate to produce carbon dioxide gas. We will investigate **what might affect the rate at which calcium carbonate reacts with acid solutions?**

Teacher modelled pitch of a closed pipe investigation

Some musical instruments use pipes to create sounds of different pitch or frequency. They are called wind instruments because the musician blows into them. We will investigate what **factors might affect the pitch of wind instruments.**

APPENDIX I

ASSESSMENT OF THE INVESTIGATIONS

Instructions for Teachers

Class TN

The assessment of this class is **teacher assessed and norm referenced**. Rank the students' work in order of the quality of their report.

Allocate
A grade for three students
B+ grade for five students
B grade for eight or nine students
C+ grade for five students
C grade for three students
Add a comment.

Students in the same group may receive different grades depending on the quality of the reports. Try to get the marked work back as soon as possible and go over class weaknesses on the whiteboard. When you give the work, back spend 20 minutes addressing problems either in a whole class setting or in a group setting, what ever you feel like. Allow students 10 minutes to write corrections in a **different coloured pen** on their work and fill in the **Improvements** section.

Class TC

The assessment of this class is **teacher assessed and criterion referenced**. Students have a copy of the marking criteria with their report sheets. Mark the students' investigations using the criterion referenced marking sheet and grade the work (A, B, C or NR) and give an overall comment. Presumably all students in a group will receive the same grade unless they submit different quality reports and obviously do not make the same contribution. Try to get the marked work back as soon as possible and go over class weaknesses on the whiteboard. When you give the work back spend 20 minutes addressing problems either in a whole class setting or in a group setting, what ever you feel like. Allow students 10 minutes to write corrections in a **different coloured pen** on their work and fill in the **Improvements** section.

Class SC

The assessment of this class is **student assessed and criterion referenced**. Students have a copy of the marking criteria with their report sheets. Hand out the master answer sheet and ask students to compare their responses with the master investigation. They should then assess their work and fill in the **Student Assessment** section. Students should determine a grade (A, B, C or NR) for each criterion. Presumably the students in a group will allocate themselves similar grades unless their reports differ. Allow 30 minutes for the assessment.

Teacher Assessed Norm Referenced Assessment Sheet for Class TN

Grade: **A** **B+** **B** **C+** **C** **C-**

Comments : (strengths and areas for improvement)

Improvements:
Write your corrections and/or improvements on your report in a different coloured pen.
Use this space to write comments that would help you to improve your performance on the next investigation.

Teacher Assessed Criterion Referenced Assessment Sheet, Class TC, and

Student Assessed Criterion Referenced Assessment Sheet, Class SC

Grade **A** if the criterion is **fully** met, **B** if **partially** met, and **C** if the criterion is not met and **NR** if the criterion is not relevant.

Marking criteria (strengths and areas for improvement) — **Comments:**

Planning

Criterion	Grade	Comments
states hypothesis	☐	______
identifies variables	☐	______
develops an overall plan	☐	______
when needed modifies plan in response to trials	☐	______

Collecting data

Criterion	Grade	Comments
describes method	☐	______
controls variables	☐	______
enters data in appropriate tables	☐	______
gathers enough data	☐	______
data is accurate	☐	______

Processing data

Criterion	Grade	Comments
presents data in an appropriate form	☐	______
relates conclusion to hypothesis	☐	______
conclusion is appropriate for data	☐	______

Evaluating the findings

Criterion	Grade	Comments
identifies sources of error	☐	______
suggests appropriate changes	☐	______

Improvements: Write your corrections and/or improvements on your report in a different coloured pen. Use space on the graph page to write comments that would help you to improve your performance on the next investigation.

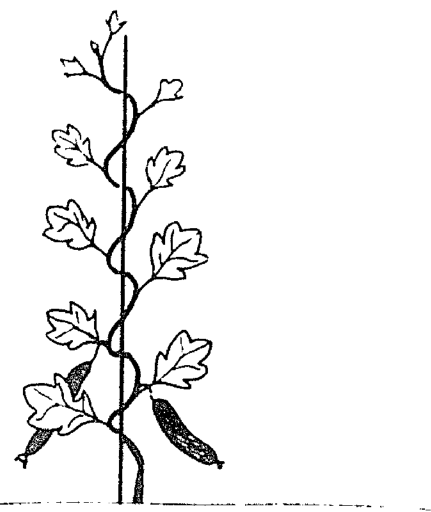
APPENDIX J

TEST OF SCIENCE INVESTIGATION SKILLS AND CODING SHEET

Test of Science Investigation Skills

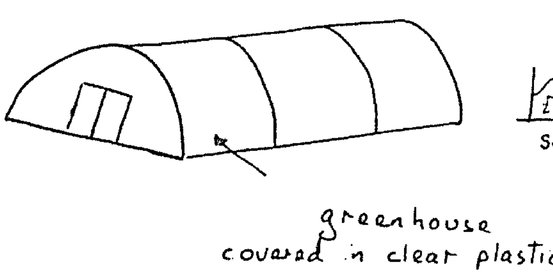
Introduction

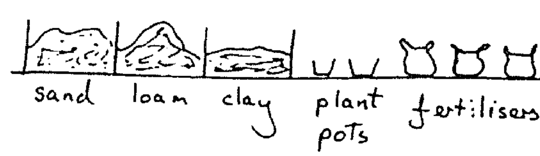
Plant breeders in the cool south-west region of Western Australia developed a new yellow cucumber which had an excellent flavour.



Consumer tests showed that people were so impressed with the yellow cucumber that they were prepared to pay twice as much for the new cucumber as they would pay for normal green cucumbers. Market gardeners were keen to grow the new cucumber as it would be highly profitable.

A horticulturist was employed to investigate the conditions under which the yellow cucumber plants would produce a large number of cucumbers. The horticulturist had a greenhouse, plant pots, three soil types (sand, loam and clay) and six different types of fertiliser that could be used in the investigation.





The horticulturist was trying to answer the question: "<u>What factors affect the number of cucumbers produced by the new yellow cucumber plants?</u>"

TOSIS:V2b

Q1. Name three factors that the horticulturist could investigate?

Q2. Write one hypothesis that the horticulturist could test in the investigation?

The horticulturist had the following fertilisers that could be used in the experiments.

Table 1

Percentage of Nutrients in Different Fertilisers

Fertiliser	Percentage of nutrients		
	Nitrate	Phosphate	Potassium
1	20	8	3
2	15	8	6
3	15	8	9
4	10	8	6
5	8	8	6
6	5	4	3

Q3. Using the materials shown in the diagram on page 1 and the fertilisers listed above, plan an experiment to test <u>the effect of nitrate concentration on the number of cucumbers produced by each plant</u>.

Describe in detail the plan of your experiment. State in your plan which of the above fertilisers you would use in your experiment.

The horticulturist had the following fertilisers that could be used in the experiments.

Table 1

Percentage of Nutrients in Different Fertilisers

Fertiliser	Percentage of nutrients		
	Nitrate	Phosphate	Potassium
1	20	8	3
2	15	8	6
3	15	8	9
4	10	8	6
5	8	8	6
6	5	4	3

The horticulturist did an experiment to test <u>the effect of potassium concentration on the number of cucumbers produced by each plant</u>.

All cucumbers were grown in a standard soil mix, one plant per pot with 30 g of fertiliser added to the soil. One plant was grown in each of fertilisers 1, 2 and 3. The number of cucumbers produced by each plant was recorded.

The horticulturist's results are presented in Table 2 below.

Table 2

Number of Cucumbers Produced by Plants
Grown in Fertilisers Containing Different Potassium Concentrations

Fertiliser	Number of cucumbers produced per plant
1	3
2	6
3	12

Q4. What conclusions can you draw from these results about <u>the effect of potassium concentration on the number of cucumbers produced by the plants</u> ?

Q5. Identify any aspect of the horticulturist's experiment that might contribute to inaccurate or misleading results and conclusions.

<u>Answer these questions by putting a circle round the best answer.</u>

Questions 6 - 8 are based on this description of an experiment on mice.

A cancer researcher was testing the effect of radiation on bone marrow tissue. One group of mice was exposed to radiation for two hours while a second group was not exposed to radiation. Both groups of mice were fed the same diet and housed in similar cages under laboratory conditions. Both groups were tested for damage to bone marrow tissue.

6. The independent (manipulated) variable is ...
   - A. amount of radiation.
   - B. diet.
   - C. type of cage.
   - D. bone marrow damage.

7. The dependent (responding) variable is ...
   - A. amount of radiation.
   - B. diet.
   - C. time.
   - D. bone marrow damage.

8. A controlled variable is ...
   - A. amount of radiation.
   - B. diet.
   - C. time.
   - D. bone marrow damage.

9. A hypothesis ...
   - A. can be proven to be true.
   - B. is a conclusion that cannot be disproved.
   - C. is a tentative explanation which can be tested.
   - D. is a theory that has been verified by other scientists.

10. Experimental results (data) ...
    - A. should be accurate and repeatable.
    - B. should be tentative (not certain) and subject to change.
    - C. should be recorded only in the laboratory.
    - D. should be recorded only outside the laboratory.

11. Which of the following is the best description of a theory?
    - A. Scientists speak of the "theory of evolution" because a theory does not have much merit.
    - B. A theory is a hypothesis that needs further testing.
    - C. Theories are hypotheses that were not supported by the results of experiments.
    - D. Theories are explanations that have been supported by the results of many experiments.

12. Scientific conclusions are ...
    - A. never amended or changed.
    - B. subject to revision.
    - C. usually unfounded.
    - D. accepted by everyone.

Test of Science Investigation Skills Coding Sheet

The total number of marks for each question are adjacent to the question number. Part marks allocated for the question are in brackets in the description of the correct answer.

Question 1: Three marks
Identifies one(1 mark), two (2 marks), or three (3 marks) relevant independent variables for testing. Acceptable variables are the soil type, fertiliser type, amount of fertiliser, amount of water, pot size, greenhouse/field, space between plants, climate, growth period, place, temperature and light.

Question 2: One mark
Writes an hypothesis as a relationship between an independent variable and the dependent variable yield/number of cucumbers (1 mark).

Question 3: Eight marks
The plan controls for phosphate and potassium, that is uses fertiliser 2, 4 and 5 (1 mark for phosphate and 1 mark for potassium).
The plan controls for amount of fertiliser (1 mark).The plan controls for soil type (1 mark).
The plan controls for sunlight (1 mark).
The plan controls for water (1 mark)
Plans to grow plants "under the same conditions" (2 marks).
The plan specifies how the dependent variable yield will be measured (1 mark).
The plan specifies an adequate sample size , that is more than one plant in each treatment (1 mark).

Question 4: Two marks
Recognises that the experimental design is poor and therefore conclusions must be tentative (1 mark).
Increased potassium concentration appears to increase yield (1 mark).

Question 5: Two marks
Recognised that nitrate concentration is not controlled (1 mark).
Recognises the sample size is inadequate (1 mark).

Question 6: One mark
Distracter (a). Identifies an independent variable from a description of an experiment (1 mark).

Question 7: One mark
Distracter (d). Identifies the dependent variable from the description of the experiment (1 mark).

Question 8: One mark
Distracter (b). Identifies a controlled variable from a description of an experiment (1 mark).

Question 9: One mark
Distracter (c). Selects a definition of an hypothesis as "a tentative explanation that can be tested" (1 mark).

Question 10: One mark
Distracter (a) Recognised that experimental results should be "accurate and repeatable" (1 mark).

Question 11: One mark
Distracter (d) Selects a definition of a theory as an "explanations that have been supported by the result of many experiments" (1 mark).

Question 12: One mark
Distracter (b) Recognises that scientific conclusions are subject to revision" (1 mark).

APPENDIX K

INTER-RATER RELIABILITY FOR THE INVESTIGATION PLANNING AND REPORT SHEET

Levels of achievement for Planning investigations, Conducting investigations, Processing data and Evaluating investigations assigned to a sample of students by two independent markers.

Student	Planning investigations	Conducting investigations	Processing data	Evaluating investigations
1	3	3	3	3
2	5(4)	3	3	3(4)
3	5(4)	3	3	0
4	4	3	0	0
5	3	0	0	0
6	6	5	4	4
7	5	4(5)	3(4)	3
8	5	4	4	4
9	5	4(3)	4	5
10	4	5	4	4(3)
11	5	3	3	0
12	4	4	3	5
13	5	4	5(4)	3(2)
14	5(4)	5	5	5
15	5(4)	0	0	0

Note 1: Where the levels of achievement differed between markers the second markers' assigned level appears in brackets.

Note 2: Levels of achievement were assigned according to the standards in Appendix C. Students who failed to complete the sections of the Investigation Report Sheet were assigned a score of zero.

Note 3: The inter-rater reliability was 0.82.

APPENDIX L

INTRA-RATER RELIABILITY FOR THE INVESTIGATION PLANNING AND REPORT SHEET

Levels of achievement for Planning investigations, Conducting investigations, Processing data and Evaluating the investigation assigned to a sample of students by two independent markers.

Student	Planning investigations	Conducting investigations	Processing data	Evaluating investigations
1	5	4	4	4
2	5	4	5	3
3	4(5)	4	4	3
4	5	5	5(4)	5
5	4	4	4	5
6	3(4)	3	2(3)	2
7	3(4)	3	0	0
8	4	3	3	4
9	3	0	0	0
10	4	3	3	3

Note 1: Where the levels differed between the first scoring and the second scoring the second scoring appears in brackets.

Note 2: Levels of achievement were assigned according to the criteria in Appendix C. Students who failed to complete the sections of the Investigation Report Sheet were assigned a score of zero.

Note 3: The intra-rater reliability was 0.88.

APPENDIX M

STUDENT QUESTIONNAIRES (CONDENSED VERSIONS)

Class TN: May

1. Look at the worksheet on the terms and words used in investigations.

Circle the number

	nothing					a lot
How much did you learn about writing hypotheses, working out the independent and dependent variables, how to control variables etc?	0	1	2	3	4	5
How much did you learn about **doing science investigations** from the worksheet?	0	1	2	3	4	5

Write down what you learned about **doing science investigations** from completing the Worksheet and give some **examples**. Be honest. If you did not learn anything about doing science investigations then you should say why.

2. Look at the lipase investigation.

Circle the number

	nothing					a lot
How much did you learn about the reaction of lipase with milk?	0	1	2	3	4	5
How much did you learn about **doing science investigations**?	0	1	2	3	4	5

Write down what you learned about **doing science investigations** from the Lipase Investigation and give some examples. Be honest. If you did not learn anything about doing science investigations then you should say why.

3. Look at the teacher modelled acid with calcium carbonate investigation

Circle the number

	nothing					a lot
How much did you learn about the reaction of acid and calcium carbonate?	0	1	2	3	4	5
How much did you learn about **doing science investigations** from this activity?	0	1	2	3	4	5

Write down what you learned about **doing science investigations** from the modelling of the acid and calcium carbonate investigation and give some examples. Be honest. If you did not learn anything about doing science investigations then you should say why.

4. Look at the hydrogen peroxide and catalyst investigation.

	Circle the number					
	nothing					a lot
How much did you learn about the reaction of hydrogen peroxide with a catalyst?	0	1	2	3	4	5
	nothing					a lot
How much did you learn about **doing science investigations** from this activity?	0	1	2	3	4	5

Write down what you learned about **doing science investigations** from the hydrogen peroxide and catalyst investigation and give examples of what you learned. Be honest. If you did not learn anything about doing science investigations then you should say why.

5. The following questions are about four activities in which you were involved.

1. The worksheet
2. The lipase investigation
3. The demonstration of acid and calcium carbonate
4. The hydrogen peroxide and catalyst investigation

Look under the heading **Planning** of an investigation.
Now think about each activity and look at them if you need.

From which activity did you learn most about **planning** investigations? ☐
Write the number of the activity in the square.
Write what you have learnt about **planning** investigations in this space and give some examples.

Look under the heading **Collecting data** of an investigation.
Now think about each activity and look at them if you need.

From which activity did you learn most about **collecting the data?** ☐
Write the number of the activity in the square.
Write what you have learnt about **collecting the data** for investigations in this space and give some examples.

Look under the heading **Processing data** an investigation.
investigation. Now think about each activity and look at them if you need.

From which activity did you learn most about **processing data** ☐
Write what you have learnt about **processing the data** for investigations in this space.

Look under the heading **Evaluating the findings** on the hydrogen peroxide and catalyst investigation. Now think about each activity and look at them if you need.

From which activity did you learn most about **evaluating the findings** ☐
Write what you learnt about **evaluating the findings** of investigations in this space.

6. These questions are about the way you learned to do investigations.
Do not use your file to answer these questions.

	Circle the number					
	not at all					a lot
How much did **talking with other students** in your group help you to learn about investigations?	0	1	2	3	4	5

Give examples of what you learnt from **talking with other students** in your group.

	Circle the number					
	not at all					a lot
How much did **talking with the teacher** help you to learn?	0	1	2	3	4	5

Give examples of what you learnt from **talking with the teacher**.

	Circle the number					
	not at all					a lot
How much did the **teacher's marking** of the hydrogen peroxide and catalyst investigation help you to learn?	0	1	2	3	4	5

Give examples of what you learnt from the **teacher's marking** of this investigation.

	Circle the number					
	not at all					a lot
How much did **making the corrections** to the hydrogen peroxide and catalyst investigation help you to learn?	0	1	2	3	4	5

Give examples of what you learnt from **making the corrections** to this investigation.

Class TC: August

1. **Look at the teacher modelled frequencies of musical notes in bottles investigations.**

	Circle the number					
	nothing					a lot
How much did you learn about the frequencies of musical notes in bottles?	0	1	2	3	4	5
	nothing					a lot
How much did you learn about **doing science investigations** from this activity?	0	1	2	3	4	5

Write down what you learned about **doing science investigations** from the teacher modelled frequencies of musical notes in bottles investigation and give some **examples**. Be honest. If you did not learn anything about doing science investigations then you should say why.

2. **Look at the investigation of the frequencies of musical notes from the ukulele.**

	Circle the number					
	nothing					a lot
How much did you learn about the frequencies of musical notes from the ukulele?	0	1	2	3	4	5
	nothing					a lot
How much did you learn about **doing science investigations** from this activity?	0	1	2	3	4	5

Write down what you learned about **doing science investigations** from the investigation of the frequencies of musical notes from the ukulele and give some **examples**. Be honest. If you did not learn anything about doing science investigations then you should say why.

3. **Look at the worksheet on marking investigations.**

	Circle the number					
	nothing					a lot
How much did you learn about **writing up investigations** from the worksheet.	0	1	2	3	4	5
	nothing					a lot
How much did you learn about **doing science investigations** from the worksheet?	0	1	2	3	4	5

Write down what you learned about **doing science investigations** from completing the worksheet and give some **examples**. Be honest. If you did not learn anything about doing science investigations then you should say why.

4. Look at the electromagnet investigation.

Circle the number

	nothing					a lot
How much did you learn about electromagnets?	0	1	2	3	4	5
How much did you learn about **doing science investigations**?	0	1	2	3	4	5

Write down what you learned about **doing science investigations** from the electromagnet investigation and give some **examples**. Be honest. If you did not learn anything about doing science investigations then you should say why.

5. The following questions are about four activities in which you were involved.

1. The teacher modelled frequencies of musical notes in bottles.
2. The investigation of the frequencies of musical notes from the ukulele.
3. The worksheet on marking investigations
4. The electromagnet investigation

Look under the heading **Planning** on your investigations.
Now think about each activity and look at them if you need.

From which activity did you learn most about **planning** investigations? ☐
Write the number of the activity in the square.
Write what you have learnt about **planning** investigations in this space and give some examples.

Look under the heading **Collecting data** on your investigations.
Now think about each activity and look at them if you need.

From which activity did you learn most about **collecting the data** ? ☐
Write the number of the activity in the square.
Write what you have learnt about **collecting the data** for investigations in this space?

Look under the heading **Processing data** investigation.
Now think about each activity and look at them if you need.

From which activity did you learn most about **processing data**? ☐
Write the number of the activity in the square.
Write what you have learnt about **processing the data** for investigations in this space.

Look under the heading **Evaluating the findings** on the investigation.
Now think about each activity and look at them if you need.

From which activity did you learn most about **evaluating the findings ?** ☐
Write the number of the activity in the square.
Write what you learnt about **evaluating the findings** of investigations in this space.

6. These questions are about the way you learned to do investigations.
Do not use your file to answer these questions.

	Circle the number					
	not at all					a lot
How much did **talking with other students** in your group help you to learn about investigations?	0	1	2	3	4	5

Give examples of what you learnt from **talking with other students** in your group.

	Circle the number					
	not at all					a lot
How much did **talking with the teacher** help you to learn?	0	1	2	3	4	5

Give examples of what you learnt from **talking with the teacher**.

	Circle the number					
	not at all					a lot
How much did the **teacher's marking** of the investigations help you to learn?	0	1	2	3	4	5

Give examples of what you learnt from the **teacher's marking** of this investigation.

	Circle the number					
	not at all					a lot
How much did **making corrections** to the investigations help you to learn?	0	1	2	3	4	5

Give examples of what you learnt from **making the corrections** to the investigations.

Class SC: November

1. Look at the electromagnet investigation.

	Circle the number					
	nothing					a lot
How much did you learn about electromagnets?	0	1	2	3	4	5
	nothing					a lot
How much did you learn about **doing science investigations**?	0	1	2	3	4	5

Write down what you learned about **doing science investigations** from the electromagnet investigation and give some **examples**. Be honest. If you did not learn anything about doing science investigations then you should say why.

2. The following questions are about the electromagnet investigation .

Look under the heading **Planning** on the electromagnet investigation. Write what you have learnt about **planning** investigations.

Look under the heading **Collecting data** on the electromagnet investigation Write what you have learnt about **collecting the data.**

Look under the heading **Processing data** on the electromagnet investigation. Write what you have learnt about **processing the data**.

Look under the heading **Evaluating the findings** on the electromagnet investigation. Write what you learnt about **evaluating the findings** .

3. These questions are about the way you learned to do investigations.

	Circle the number					
	not at all					a lot
How much did **talking with other students** in your group help you to learn about investigations?	0	1	2	3	4	5

Give examples of what you learnt from **talking with other students** in your group.

	Circle the number					
	not at all					a lot
How much did **talking with the teacher** help you to learn?	0	1	2	3	4	5

Give examples of what you learnt from **talking with the teacher**.

	Circle the number					
	not at all					a lot
How much did **marking your own** investigation help you to learn?	0	1	2	3	4	5

Give examples of what you learnt from **your marking** of the investigation.

	Circle the number					
	not at all					a lot
How much did **making the corrections** to the investigations help you to learn?	0	1	2	3	4	5

Give examples of what you learnt from **making the corrections** to the investigations.

4. These questions are about all of the investigations.
You will need to **go back to your file** to refresh your memory.

	Circle the number					
	not at all helpful					very helpful
How effective were the **sheets on investigations** in **helping you learn** to do investigations?	0	1	2	3	4	5

Give examples of what you learnt from **completing the sheets on investigations**.

	Circle the number					
	not at all					a lot
How much do you think that you **have improved at doing** investigations?	0	1	2	3	4	5

Give examples of improvements you have made.

	Circle the number					
	not at all					a lot
How much do you think that you **have improved at marking** your own investigations.	0	1	2	3	4	5

Give examples of what you learnt from **marking** your investigations.

5. Seven different **ways of learning** how to do investigations have been listed below. Which is the **best way for you to learn** how to do investigations? Rank the ways of learning from 1 to 7 with number 1 the best and number 7 the worst.

	number from 1 to 7
watching and listening to the teacher model an investigation	☐
watching and listening other students do an investigation	☐
talking with the teacher	☐
talking with other students	☐
doing an investigation by yourself	☐
doing an investigation in a group	☐
marking the investigation	☐

6. In all of your investigations your group marked their own work.

Circle which you prefer.

group marking or teacher marking

Give reasons for your answer.

7. Write any other comments that you would like to make about the investigations here.