Science and Mathematics Education Centre

Improvement of students’ scientific writing in a middle-years Science classroom

Duncan Wood

This thesis is presented for the Degree of
Master of Philosophy
of
Curtin University

September 2015

DECLARATION

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgment has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Signature:………………………………………………

Date: 2/9/15

ABSTRACT

The aim of this study was to investigate the effectiveness of scaffolding, modelling, peer/self-assessment and metacognition in promoting the writing skills of Year 6 Science students. It focussed on students' ability to write reports using data gathered from science experimentation in class or from secondary sources of data. Changes in their attitude to writing, self-efficacy and actual performance in the writing tasks were tracked and problems with the improvement processes identified.

Research evidence suggests a tendency for many students to find scientific writing assessments challenging due to factors such as the differences in style and requirements from other subjects and vocabulary. The students in the study showed a range of initial abilities given their relative inexperience with writing in science. Participating students actively engaged in the programme which resulted in improved scientific writing skills. This improvement was further enhanced by regular formative assessment tasks built into the intervention process. Due to the fact that students were often assessed on scientific writing, as a result their grades also improved.

Drawing from the existing literature (Keys, Hand, Prain & Collins, 1999), this project utilised improvement strategies such as a scaffold for writing with strong and weak model answers. Self and peer feedback (Butler & Nesbit, 2008; Rice, 1998) of student writing and metacognitive reflection about effective writing strategies were also used. Attitude and self-efficacy were measured by the use of interviews with eight students before and after the use of the strategies and in student journals kept by the entire class. A diary kept by the teacher during the implementation period was the major instrument for describing perceived flaws in the implementation of the improvement strategies. These qualitative sources of data were analysed thematically to identify the effectiveness of the programme. The improvement in the students' scientific writing was primarily measured by work samples which were assessed to provide quantitative data. This mixed methods approach meant that findings were supported by data from numerous different sources, both qualitative and quantitative. No one source told the entire story and the claims that have been made by the researcher had several supporting pieces of evidence.

The writing improvement strategies were found to improve students' views about, and their self-efficacy towards, scientific writing. The students experienced success with writing and saw improvements with sustained effort. As a result, students' attitude to writing and their view of their ability to tackle writing tasks was much more positive. Students writing scores in their practice of scientific conclusions increased significantly which led to improved grades at the end of the semester. A major limitation in the use of the improvement strategies was the extra time requirement imposed on the class. Considerable time was required for the training so that students were able use the strategies effectively.

ACKNOWLEDGEMENT

I would like to thank Dr Rekha Koul and A/Professor Bill Atweh for the considerable time spent in the planning, implementation and writing of this research and thesis. Their advice and knowledge has been invaluable.

I would also like to thank my wife and family for their love and support. Their encouragement made all the difference.

TABLE OF CONTENTS

TABLE OF CONTENTS - CONTINUED

TABLE OF CONTENTS - CONTINUED

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1: OVERVIEW

1.1 Introduction

Chapter one provides an introduction to the purpose and background of the research. It is divided into five sections with the next section (1.2) describing the background to the project. Section 1.3 discusses the nature of the research problem. Section 1.4 outlines the rationale for the research project and its significance. Section 1.5 deals with the specific research questions and Section 1.6 describes the research paradigm and methodology for the research. Section 1.7 outlines the structure of the thesis and Section 1.8 summarises this chapter.

1.2 Background

It is widely accepted that students' literacy levels directly affects their ability to access learning in today's educational institutions. The ability to read and comprehend information and instructions influences how a student can acquire and understand knowledge presented to them in a classroom. Moreover, a students are often asked to present their knowledge and understanding in a piece of writing for both formal and informal assessment. Teachers must therefore be cognisant of the fact that a particular student's ability with written communication can influence his/her performance in a subject area.

This research was centred on the improvement of students' scientific writing in a Victorian private school. In this thesis the term 'scientific writing' is used to describe the pen-and-paper pieces of writing that students perform in secondary school science classes. One of the more common forms of scientific writing is the conclusion of an experiment. This is often used to gauge the depth of understanding of concepts from the results of experimentation for both primary and secondary data. In particular, the writing of a scientific conclusion allows students to convey their understanding of what has happened in an experiment and (perhaps more importantly) why it happened in that way. This insight into the student's thought processes requires a certain level of communicative ability. The student should be able to put their thoughts coherently and logically into a paragraph or more of writing. Baker et al. (2010, p.105) go further and suggest that “writing promotes critical-thinking skills and construction of vital scientific concepts and challenges ingrained misconceptions”. It is without a doubt that scientific writing is an important part of the Science classroom for both the students’ learning and the ability of teachers to assess that learning.

It follows, then, that Science teachers have a responsibility to teach students how to structure and improve their writing in a scientific context in addition to the more conventional teaching areas of science theory, language and experimental practice. Sometimes this can prove problematic, when the teaching of ‘writing’ is thought of as the domain of the English language teacher and can be underemphasised in the Science classroom.

1.3 Problem

In many subject areas, there is technical language or vocabulary that is specific to the subject. There often are conventions of writing that determine a style in which a piece is written. Secondary school Science, in particular, has many terms and phrases that are specific to the discipline or have entirely different meanings than those in other subject areas. A common writing piece is the experimental report for which students are asked to present and explain the findings of an experiment using logical clauses and making links to scientific theory or law. These can be major stumbling blocks for students who find writing difficult or for those with inexperience with the genre. Students with poor writing skills or who are new to science may not understand the requirements of the writing that they are asked to complete. This can result in low achievement scores on writing tasks and be linked to motivational issues for those students. The problem, then, is how best to expedite the learning of the genre or to increase the level of students’ existing science writing skills.

As a Science teacher, the researcher realises that the skill of writing scientific discourse can be improved dramatically for novice science learners and the problem becomes one of discovering effective strategies to realise this. These strategies can be implemented as ‘interventions’ in a classroom where there are students who have difficulties with writing or inexperience with writing in the scientific domain.

1.4 Rationale and Significance

At the time of the research, the author was a teacher of middle-school science at a K-12 college in Victoria. The project coincided with the authors desire to get the best outcomes for his students. For a participant researcher, the research serves to improve the author's ability to address the needs of his students in the area of scientific writing. This seemed to be an area in need of considerable improvement, as little class time was assigned to addressing scientific writing skills. Because scientific writing constituted a large proportion of the formal assessment of the researcher's students it seemed appropriate that this was an area that needed more focus. As the classroom teacher, being able to assess the worth and observe the effect of teaching strategies that address writing skills was considered important and worthwhile. The researcher stood to gain a great deal of personal development in terms of the teaching methods and a more in-depth understanding of the relative worth of the teaching methods employed.

The participating students also stood to benefit from the research as it could significantly benefit their learning and attitude towards science and scientific writing. Students who experience success and are cognisant of the journey they have taken to improve will feel better about themselves. They will be able to better communicate their ideas and understandings about what they are learning, therefore broadening the way in which they are able to learn about science and apply scientific reasoning to various contexts. As they are required to submit formal assessments which use scientific writing as a key component, it is likely that they will perform better in these assessments after the research project has been completed and that they will, in turn, receive better grades. As they become more successful in science, their interest in the subject may increase, which could influence their subject choices and even career choices in future. The students may also start to use the techniques that they learn during the study in other areas of their schooling.

The research school's Science Department had created a series of writing scaffolds that were presented to the students as guidelines for writing about experiments. They were often presented as acronyms to help students to remember what was important to include in the writing. One of the acronyms, F.R.E.S.H. (Further experimentation, Reliability, Evidence, Science theory, Hypothesis), was used to remember the key components of a conclusion to an experiment (and this is further elaborated on in Chapter 2 and 3). The research study would also serve to validate the use of this scaffold as an effective teaching strategy.

The participating school would also stand to benefit as if would be able to use the outcomes of this study to aide in the professional development of its science teachers in the area of science writing. This could result in a paradigm shift in the way science is taught at the school, influencing the time allocated to the teaching of scientific writing in the curriculum.

In addition to those directly affected by the research, this thesis will also serve as a validation of these strategies that other researchers in this field, such as Rutherford (2007), Hand, Wallace and Yang, (2004), LaConte and Berry (2006), Brandt (1971), Conners (2007), Butler and Nesbit (2008), Akerson and Young (2005), Rice (1998) and Butler and Nesbit (2008), have found to be effective. Other teachers may be more inclined to use them as a result.

1.5 Purpose and Research Questions

The purpose of the research was to measure the changes in learning, student attitudes and self-efficacy towards scientific writing as a result of implementation of intervention strategies (which are elaborated on in Chapter 3). In order to achieve this, the study focused on four research questions:

Research questions:

1) During the implementation of the intervention strategies, are there changes in students':
    a) Views of themselves as scientific writers?
    b) Self-efficacy towards scientific writing?
    c) Performance in scientific writing?
2) What problems arise from the implementation of the intervention strategies?

1.6 Research Paradigm

According to Willis (2007, p. 8), a research paradigm is "a comprehensive belief system, world view, or framework that guides research and practice in a field". In educational research, this means a framework of ideas that governs how one approaches the research. This can affect the type of data collected, its analysis and in which it will be analysed, presented and used. The framework that governs the researcher's world view is a post-positivist/interpretive one in which more qualitative data is collected (e.g. case-studies and interviews) in order to either disprove a hypothesis or support it while allowing that it may not be proved without a doubt.

Ontology is the way in which one views reality – what is real or perception of what is real. There are ways of viewing what "exists". One person might say reality is the physical world (all that there 'is' in terms of matter), whereas another may say that the way in which you perceive things is creating reality in your mind. A researcher's view of reality could perceivably influence the paradigm selected for research. The researcher, in this case, has the view that reality is what is perceived by the individual. This is a mental ontology which is reflected in the way thoughts and comments of the research subjects have been gathered, as well as the researcher's own feelings being part of the study. The researcher sees perception as a result of the subjects' prior experiences and how they apply that to events and their current emotional state.

Epistemology is the way you view 'knowledge' of reality – what knowledge is, how you gain this knowledge, and ways in which this knowledge can be affected by others. From ontology comes the question of how a person finds out about the reality that they have accepted, (i.e. gaining knowledge and the limits of that knowledge). Their epistemology comes from the act of asking these questions. A material (physical) ontology would suggest that knowledge can be found by experimentation and interaction with the physical reality (empirical epistemology). However if one accepts that reality is merely perception and could be manipulated somehow (a more mental ontology) then one must accept that knowledge is subject to limitations or changes to that perception imposed by others or the environment. This second statement resembles the epistemology of the researcher, namely that perceptions change and so knowledge changes. Mulholland and Wallace (2003, pp. 881-882) believe that their perception (and therefore knowledge of reality) could be changed by their own experiences during research, stating that "…we call on our own experiences, knowledge and theoretical dispositions, to collect data and present our understandings". This means that knowledge is value laden and requires analysis over time, bearing in mind that the researchers' perceptions will also evolve. In educational research, the manner in which we collect data may be affected by the epistemology of the researcher in terms of the level of scrutiny given to reliability or external factors such as perceived gender or cultural bias. Some of these concerns are be addressed in the ethics section (3.7) of Chapter 3.

Methodology is the overarching idea(s) behind the way you collect, analyse and present data as informed by the ontology and epistemology of the researcher. Methodology in this context is an overarching way of conducting research. The researcher chose a mixed-methods approach for this research that involves multiple case studies, including observations, interviews and reflective journals to expose the longitudinal development and perceptions of eight of the researcher's students in detail. A further 11 students were examined utilising journals and quantitative results. The researcher also contributed with reflections during the journey through an intervention process aimed at improving student writing in a middle-years science classroom. The researcher chose to use this methodology as he used a constructivist epistemology, wanting to participate and narrate the evolution of perceptions of others over time:

> "narrative researchers describe such lives, collect and tell stories of them, and write narratives of experience". (Connelly & Clandinin, 1990, p. 2)

Using an interpretive paradigm to shape the research, the researcher sought to tell the tale of some of the changing perspectives of participants in the research. Such a research problem requires the uses of longitudinal case-studies (in the form of interviews before and after the study) of the participants. Realising that the researcher's own experiences and perceptions would be changed as a result of this process, the researcher also shared his own case-study so as to see the interpretation of the results as a "joint construction by the participants and myself as author" as Lloyd (2007, p. 67) puts it. This is combined with more qualitative data in the form of work samples and journaling and also quantitative data in the form of numerical scores that show improvement – hence ‘mixed methods’.

A mixed methods approach allows triangulation of the data in order to answer the research questions or potentially inform further quantitative or qualitative research. The nature of the research questions involves the perceptions of students but also their improvement which is readily quantifiable:

> “Quantitative data, such as scores on instruments, yield specific numbers that can be statistically analysed, can produce results to assess the frequency and magnitude of trends, and can produce useful information if you need to describe trends about a large number of people. However, qualitative data, such as open-ended interviews that provide actual words of people in the study, offer many different perspectives on the study topic and provide a complex picture of the situation.” (Cresswell, 2005, p. 552).

The other advantages of this approach are that it better taps the perceptions of both the researcher and subject(s) alike. It gives a more ‘human’ feel to the research. Changes in values and reflective practises can be mapped over time, thus allowing links to be made between experience and perception as long as the interviews are of sufficient quality. It allows the use of qualitative data which can possibly be more informative to the researcher given an appropriate context.

Some limitations to this paradigm include the dependability of the sources which require the use of an audit trail. A sense of dependability by frequently using direct quotes from the interviews allows the reader to gauge the authenticity of the researcher’s conclusions. Another limitation is that of transferability, which involves the available case studies being diverse enough to allow the reader to make sense of the findings in a different (even their own) context.

1.7 Outline of Thesis

The thesis is divided into five chapters with the following themes. Chapter 1 provides an overview of the research and the research problem. The second chapter is devoted to a review of the current literature on the subject of scientific writing and how to improve it. Chapter 3 deals with the methodology for the research project and provides a context for the participants. Chapter 4 reports the qualitative findings from interviews and diary entries and the quantitative findings from the students work samples. The fifth chapter provides conclusions and implications of the research.

1.8 Summary

This chapter has given an overview of the background and rationale for the research project. Section 1.2 gave background information about the use of scientific writing in secondary schools. It identified scientific writing as an underemphasised but important part of science education as well as describing its role as an assessment tool for student understanding in science education. In Section 1.3, the research problem of how best to improve the scientific writing of students in a secondary school was outlined. Section 1.4 discussed the significance of the study, identifying benefits to many of the participants and as a validation of methods outlined by other researchers. Section 1.5 gave a description of the research questions that the study intended to answer. In Section 1.6, the research paradigm, ontology, epistemology and methodology were described and the final section (1.7) provided an overview of the other chapters in the thesis.

The next chapter provides a review of the relevant literature for the study. It gives a rationale for the use of scientific writing as an assessment tool, as well as describing the nature of the intervention strategies utilised during the research project.

CHAPTER 2: REVIEW OF THE LITERATURE

2.1 Introduction

A comprehensive review of the literature relevant to the research is presented in this chapter. Section 2.2 provides a brief explanation of the action research paradigm used in this research. Section 2.3 engages with the need for writing in science as a communication tool and its use to inform the educator from an assessment perspective. The next three sections elaborate on the types of learning strategies that help students to write more effectively and were used in this research, namely: the use of Feedback, Metacognition, and Scaffolding. The final section provides a conclusion and briefly summarises the key issues discussed in the literature review.

2.2 The Action Research Paradigm

This research sought to change the scientific writing skills of students in a Middle School classroom. A consequence of the research is that fellow educators could be challenged to consider their own students' learning in this regard. Furthermore, this could result in a change to their teaching practices if needed. In an editorial on action research Kyle (1997) says that "Action research is part of a long tradition among grassroots activists to seek mutually constitutive relationships between research and social change. Action research differs from traditional empirical–analytic and interpretive research in both its dynamism and its continuity with an emergent practice." It is clear to the researcher that this project can be categorised as falling under the action research paradigm.

The appeal to the teacher-researcher as a core part of the research project is summarised eloquently by Kayaoglu (2015) who conducted research on the viability of action research as a potential model for professional development.

> "Action research, which places the teacher in the core of professional development, is quite new in the sense that it poses a radical change to the heavily centralized education system. This approach therefore has the potential

> not only to revolutionize the way in which the teacher is perceived in relation to their new competences but also to inspire a paradigmatic change in teacher education since it fundamentally moves away from the prevalent model of teacher-centred education characterized by a heavy focus on conventional theories of teaching/learning. Teachers are empowered to investigate what actually happens in the classroom with a vision of challenging and changing it when/where pedagogically necessary.” (Kayaoglu, 2015, p. 140)

The empowerment felt by the teacher-researcher engaged in the investigation, intervention and perceived solution to an issue for students in his own classroom was undoubtedly a powerful motivator and impetus for the research project.

2.3 Purpose of Writing in Science Education

The researcher believes that science education is designed to produce students who are informed and knowledgeable about the universe in which they live. It encourages students to question why things happen, and how they occur. In order to find answers to these lofty questions, students must be able to research, experiment and communicate ideas which require a high level of scientific literacy.

> “In addition to its practical applications, learning science is a valuable pursuit in its own right. Students can experience the joy of scientific discovery and nurture their natural curiosity about the world around them. In doing this, they develop critical and creative thinking skills and challenge themselves to identify questions and draw evidence-based conclusions using scientific methods. The wider benefits of this ‘scientific literacy’ are well established, including giving students the capability to investigate the natural world and changes made to it through human activity.” (ACARA, 2016, Accessed 14/2/2016)

Norris and Phillips (2003) describe the idea of ‘scientific literacy’ as having two senses: A fundamental one of “reading and writing when the content is science” (Norris & Phillips, 2003, p. 224) and a derived one for which you are “knowledgeable, learned, and educated in science” (Norris & Phillips, 2003, p. 224). They make these definitions to show the link between ‘English literacy’ and ‘Science literacy’ but also to show how in a text-dominated discipline like Science, you can’t have one without the other.

In order to adequately engage with the ‘text of Science’, students must have a large technical vocabulary. For some, this is quite a hurdle:

> “Students must traverse an intricate linguistic landscape in which they meet interlocking definitions in texts packed with novel, technical terms that are difficult to pronounce.” (Hohenshell, Woller & Wallace, 2013, p. 38)

The vocabulary increases in size and complexity as a student progresses through their schooling. However just having the vocabulary is not enough. Often, this language must be communicated in a logical and meaningful way, taking the form of a scientific report on an experiment which students have conducted, for example.

‘Science’ is not just a set of facts and figures to be remembered and understood. Students need to be able to apply these concepts in the realm of scientific experimentation. They must also learn the skills of planning, predicting, experimentation, observation, interpretation and drawing a conclusion. (ACARA, 2016, Accessed 14/2/2016). The Science teacher recognises that these skills need to be assessed on a regular basis to better gauge their students’ progress in these areas. Some of the more common ways of assessing these skills include observation of practical investigations and short answer tests (Shepardson & Britsch, 2001). However, the conclusion to an investigation is often assessed as a written task that allows a student to communicate his or her interpretation of the data and the scientific reasoning that supports it. The rationale for this is that this kind of response is open-ended and allows for the identification of deeper understanding.

Shepardson and Britsch (2001) provide an overview of various types of Science assessment that could be performed in lieu of or alongside more traditional quizzes and tests. They elucidate forms of ‘profiling’ assessment such as practical tasks, open responses, peer/self-assessment, graphic products, observational checklists, child interviews and journaling. Shepardson and Britsch (2001) describe scientific writing as “paper-pencil tasks that provide children with information and ask children to apply their understandings to the situation described in the task. They do not ask children to restate the facts; they allow children to respond in a variety of ways to explain and apply their ideas” (p. 124). Scientific writing in this manner also allows the teacher to more readily discover misconceptions as it may give evidence of thought processes and show the foundations of the current knowledge. Arends (2001) stated “Many teachers and test-experts agree that essay tests do the best job of tapping students’ higher-level thought processes and creativity” (p. 213). It is important to note that many researchers have different names for writing assessments, but they are pointing to the same practice of writing about scientific concepts or experiments.

The researcher values the open-ended nature of scientific writing assessments which can accurately pinpoint the level of a student’s scientific reasoning skill. This is because the researcher believes that education should be developmental, and new knowledge should not be placed over shaky foundations. Klassen (2006) acknowledges the evolutionary trend of assessment in science over the past 25 years of adapting to a more student centred (constructivist) approach to Science education. Assessments in science need to be open ended enough to allow for entry by learners at different stages of their learning, while also trying to elicit a detailed description of thought processes. It is assumed that scientific writing tasks can fill this role as they do not have an immediate ‘right’ or ‘wrong’ answer.

Moreover, the need for written assessments is eloquently put by Krajcik and Sutherland (2010): “literacy practice [is] essential to fostering inquiry in the classroom… engaging students in constructing explanations and arguments, which are essential components of scientific discourse” (p. 458). Krajcik and Sutherland (2010) suggest that there are five ‘aspects of literacy’ important to science inquiry:

> “(i) linking new ideas to prior knowledge and experiences, (ii) anchoring learning in questions that are meaningful in the lives of students, (iii) connecting multiple representations, (iv) providing opportunities for students to use science ideas, and (v) supporting students’ engagement with the discourses of science” (Krajcik & Sutherland 2010, pp. 456-457)

Some teachers may resent this practice as ‘literacy’ requires explicit teaching within the course which could detract from time spent teaching practical skills or science concepts. Baker, Barstack, Clark, Hull, Goodman, Kook, Kraft, Ramakrishna, Roberts, Shaw, Weaver, and Lang (2008, p. 106) suggested that many middle school science teachers were hesitant to include “writing-to-learn” activities due to a lack of familiarity with them and therefore being less comfortable with them. However, most science curricula have ‘communication’ as a specific or overarching outcome which needs to be addressed. For example, the Australian Curriculum states that Communication is one of the key areas in Scientific Inquiry along with Questioning and Predicting, Planning and Conducting, Processing and Analysing Data and Information, and Evaluating (Australian Curriculum, 2015, Accessed 23/2/15).

An example of scientific writing is the conclusion written at the end of a laboratory report. A scientific conclusion is often presented as an argumentative statement which must be supported by links to supporting data or examples and relevant scientific theory. A degree of reasoning is required and this needs to be conveyed through the writing. Wollman-Bonilla (2000) said: “Although there is more to doing science than its language, scientific reasoning is a linguistic process” (p. 37). It is clear that the use of scientific writing both as a learning and assessment tool has many benefits. It gives an insight into the thought processes students are making during scientific inquiry. By writing, students can show a logical rationale for their evaluation of first or second hand data resulting from experimentation that may be challenging to obtain from other forms of assessment.

It has been shown that students will perform quite differently on these scientific writing assessments depending on their different backgrounds and proficiencies with similar forms of assessment. In a study performed by Lawrenz, Huffman and Welch (2000), nearly 3,500 grade 9 students in the United States were tested using varied styles of assessment including a “written open-ended test”. The aims of their study were to identify differences in performance on different assessment items based on race or gender differences. Caucasian students were found to outperform others in scientific writing assessments. They speculated that this was due to language difficulties of students coming from non-English speaking backgrounds. This raises obvious questions about fairness and reliability of these assessments with respect to cultural background. It is also conceivable that students whose scientific background put a lot of emphasis on rote responses would not fare as well on a more creative/open-ended style of assessment. They also found that "the student outcomes on the different assessment formats are more highly correlated for higher achieving students than for lower achieving students" (Lawrenz, Huffman & Welch, 2000, p. 279), meaning that students who did well on multiple-choice tests were also likely to do well in writing tasks and hands-on activities, whereas the low achieving students might do better on some tasks but not others.

Lee and Luykx (2005) summarised the work of Ruiz-Primo and Shavelson (1996) and Shaw (1997) surrounding scientific writing assessments where students respond freely to experimental results or stimuli saying "performance assessment tends to rely heavily on students' ability to read and write, confounding literacy skills with content knowledge" (p. 419). Furthermore, Lawrenz, Huffman and Welch (2000) postulate that lower achieving students did better on hands-on assessments where the teacher assessed them using observation as they were perceived to be more 'fun' or 'different' activities. By implication, then, a scientific writing assessment, such as a concluding an investigation, would be considered 'boring'. Lawrenz, Huffman and Welch (2000) also said that the "hands-on assessment formats may require less cognitive ability because the situations are less abstract than ones posed in written formats where students have to imagine the materials and results" (p. 288). Therefore, the gap between the low and high achievers would widen as the low achievers 'switch off' if the written format is too hard to understand. This also impacts on the students' view of themselves as scientists and science writers.

Bandura (1994) notes that students who doubt their capabilities find it hard to commit to difficult tasks and often fail because they think that they will. He said: "Perceived self-efficacy is defined as people's beliefs about their capabilities to produce designated levels of performance that exercise influence over events that affect their lives" (Bandura, 1994, p. 71). It is therefore important to acknowledge that the students' mindset when entering a written assessment on scientific conclusion is going to affect their performance. Bandura believes that a strong sense of self-efficacy can be fostered by allowing students to achieve success in 'mastery experiences'. In my research, the mastery experiences involved the use of the students' scientific reasoning within scientific writing exercises where they reflect on achievement of increasing complex criteria. Special considerations need to be made for students with weak literacy skills, cultural differences and low self-efficacy so that the writing tasks are accessible to all who attempt them – a ‘one size fits all’ approach would not provide the mastery experiences required for strengthening self-efficacy:

> “Those who have a high sense of efficacy visualize success scenarios that provide positive guides and supports for performance. Those who doubt their efficacy visualize failure scenarios and dwell on the many things that can go wrong. It is difficult to achieve much while fighting self-doubt.” (Bandura 1993, p. 118)

Given the importance of science writing in developing student understandings and its use as an assessment item, Baker et al. (2008, p. 107) suggest that a variety of strategies may be useful for increasing the skill level of students in scientific writing exercises in a Science classroom including: Teacher modelling, writing heuristics, written feedback, peer/ self-assessment, face-to-face conferences and group discussions of writing samples. Akerson and Young (2005, p. 38) applaud the use of journals with daily writing prompts as a way of developing students’ “informational writing skills” and encouraging their students to link observations to inferences – the essence of writing a conclusion.

My research project centred on the use of three main strategies - feedback, metacognition and scaffolding/modelling. There has been a great deal of research done in these three areas, as is elaborated on in Sections 2.3, 2.4 and 2.5.

2.4 Writing Learning Strategies - Feedback

From the researcher’s point of view feedback is the response to a piece of student work given by their teacher, peers or themselves. Constructive feedback is that which provides the student with a way to improve on what they have done. Statements such as ‘good job’ would not be considered constructive. Hattie (1992) performed a synthesis of 134 meta-analyses of “educational innovations” designed to improve student outcomes and suggested the largest effect size on achievement is feedback:

> "An effect-size provides a common expression of the magnitude of study outcomes for all types of outcome variables, such as school achievement. An effect-size of 1.0 indicates an increase of one standard deviation, typically associated with advancing children's achievement by one year, improving the rate of learning by 50%, or a correlation between some variable (e.g. amount of homework) and achievement of approximately .50. When implementing a new program, an effect-size of 1.0 would mean that approximately 95% of outcomes positively enhance achievement, or average students receiving that treatment would exceed 84% of students not receiving that treatment." (Hattie, 1992, pp. 5-6)

Hattie found that the effect-size in standard deviations for "reinforcement [was] 1.13, remediation and feedback .65, [and] mastery learning (which is based on feedback) .50" (1992, p. 9). Given the implications of this study, it would seem remiss for any intervention into students' writing not to include feedback in some form as a part of the planned improvement process.

Feedback is deemed by Brandt (1971), Rice (1998) and Butler and Nesbit (2008) to be an essential part of improving students' writing. In this research project, emphasis was placed on peer/self-assessment (and hence feedback) rather than teacher feedback. As the students themselves became the providers of feedback, an extra dimension was added to their contact with scientific writing. They had contact with more than just their own work, and learned from the experience of giving feedback.

Brandt (1971) suggested that, when students undertook critical reading of flawed pieces provided by the teacher, they improved their own writing. He advocated the introduction of flaws that would concentrate on "the coherence of a piece of writing and the clarity of the interpretation" but suggested that the flaws would need to be carefully added so that a "careful but inexperienced" student would be able to participate and still get something from the activity. The students are able to make links between what makes a piece 'good' and their own work in terms of coherence, logic and requirements necessary for success. They also make links between common flaws and occurrences of these in their own writing. The researcher also sees this as a ‘training’ opportunity for students whose experience with assessing others’ work is little to none. This would become a key activity for the students learning to peer, and self-assess their work.

Continuing the vein of peer assessment, Rice (1998) asked the students themselves to distribute their own work to receive constructive criticism from their peers. In his university level ‘Science writing’ course, one of the main classroom activities to improve the students’ writing was to conduct peer-assessment. Rice notes the benefits of peer assessment in the following quote:

> “As they become more proficient in picking out writing problems and suggesting alternative strategies for others, they can turn these skills to their own writing” (Rice, 1998, p. 270)

There are pitfalls with this strategy as Rice (1998, p. 270) notes, “While a few students enjoy this from the outset, most are initially reluctant to criticize other papers because, as many of them tell me, they do not feel competent to point out problems in someone else’s writing”. Of course, the proficiency comes with practice and these reluctant feelings fade, but it seems obvious that a guiding hand would be required for the much younger students engaged in this research project who suffer from inexperience and a lack of maturity that the university students would have in Rice’s study. The clash between students’ ability to assess each other in a constructive and genuine way and their desire to be accepted socially can have implications for the effectiveness of peer-assessment. As Snowman and Biehler (2000, p. 90) point out, “the desire to conform reaches a peak during the middle school years” and this can alter their opinions. They then suggest it may be preferential to “invite them to write their opinions anonymously”. The ‘guiding hand’ mentioned earlier might also be realised in the use of rubrics to guide the student’s assessment of each other’s work:

> “A rubric describes various levels of performance that students are supposed to attain across a scoring scale. A rubric is therefore important for both teachers and students as it reveals the desired achievement for a performance task with an established set of criteria so as to score students' performances.” (Kocakülah, 2010, p. 146)

The design of success criteria or a marking rubric with input from the teacher or other members of the class has been shown to be effective in giving students a way to write feedback in a more constructive way that removes the social/emotional aspects of the feedback that the younger students may be inclined to write. While Butler and Nesbit (2008) also described the role of peer and teacher feedback in various formats, they suggested that self-assessment was an important process for the improving writer. During their research they implemented a schedule of self-feedback sessions that utilised a rubric developed in partnership with the teacher and the student. The rubric focused on "the clarity of writing and the accuracy and completeness of the scientific content" (Butler & Nesbit, 2008, p. 139). The criteria helped the students to review their work. The next phase was peer feedback which concentrated on clarity. Students were encouraged to ask themselves if the passage communicated the student's thoughts effectively. Butler and Nesbit (2008) also said that written feedback from the teacher had to address specific writing skills and scientific knowledge rather than spelling and labelling or giving a percentage score. It was suggested that the feedback be given separately to the written piece to encourage ownership of the written work. Individual conferences with students could be held to provide feedback for specific and persistent misconceptions. The authentic and motivating experience of publication was outlined to encourage students to review their work for sentence structure, clarity and science content (Butler & Nesbit, 2008).

In all of the cases above, students who gave and received feedback became better scientific writers

Apart from feedback given from a marking rubric/success criteria, written feedback in terms of general comments do have their uses (as long as they are free from social/emotional phrases or 'pat-on-the-back' style comments). Furthering the analysis of Brandt's (1971) work on the introduction of artificially flawed pieces of scientific writing, he insists that students must identify the flaw and correct it, saying why it was flawed. Types of flaws include: statements which do not match the results, statements which are vague and noncommittal, gaps in reasoning or missing components. This teaches students about constructive criticism and also about being objective when addressing flaws. As the piece is artificial, students felt there were no social ‘restrictions’ on the level or quantity of criticism they could make – the owners ‘feelings’ were not an issue.

Rice (1998) describes the criticism of writing as having two aspects - local and global. Local criticism occurs when one addresses the technical aspects of writing like grammar and sentence length which are right or wrong according to conventions. Global criticism occurs when one addresses the overall organisation and ‘fulfilment of expectations’ of the document. With this in mind Rice has his students critique each other’s scientific writing before it is submitted and allows time for review (Rice, 1998).

2.5 Writing Learning Strategies – Metacognition

Arends (2001, p. 468) says that “metacognition refers to learners’ thinking about their own thinking and their abilities to use particular learning strategies appropriately.” He also communicates that it is important that students identify which strategies are best and when to use them for themselves. These strategies might be developed independently, but it is thought that, if a few strategies are taught, the students would be more likely to choose one appropriate to their own learning style.

Metacognitive knowledge is able to be developed according to Flavell (1979, p. 907) who notes that “metacognitive knowledge consists primarily of knowledge or beliefs about what factors or variables act and interact in what ways to affect the course and outcome of cognitive enterprises. There are three major categories of these factors or variables—person, task, and strategy.” Flavell (1979) describes personal metacognition as awareness of your own cognitive processes or those of others. He considers task metacognition to be an understanding of how a particular task will affect cognitive processes. Finally, Flavell (1979) describes strategy metacognition as the use of particular learning or problem solving strategies to help reach a particular goal.

Metacognitive students have been shown by Conners (2007) to be better writers when they identify productive and non-productive strategies that they use when performing scientific writing. They understood that planning and time management were important parts of the writing process amongst other things. Conners (2007) used a journal with predefined prompts such as "Something I learned today…" and "I'm stuck on…" (Conners, 2007, p. 3) to gauge the amount of self-questioning that the student performed after each class (with regular collection). Interviews with students, as well as the journal and other observations and records of student work, were used to ascertain the awareness and use of metacognitive strategies that students employed during essay writing. The strategies were planning (essay skeleton/structuring use of mnemonics, etc.), monitoring (time management/ organisation/information management strategies like 'trash and treasure', etc.) and self-questioning (reflections on knowledge and work practices). Conners was able to show a trend between high use of metacognitive strategies and high performance on scientific writing activities.

Butler and Nesbit (2008) and Akerson and Young (2005, p. 38) advocate the use of science notebooks as a "record of scientific experiences" (Butler & Nesbit, 2008, p. 137) but more importantly involving "students in the process of constructing knowledge" (Butler & Nesbit, 2008, p. 137). After writing about an experiment, students were encouraged to draw a "line of learning" (Butler and Nesbit, 2008, p. 138) which was defined as being when "Students develop a deeper understanding of the target concept. They apply the concept to new situations and learn new science vocabulary." Here, scientific vocabulary could be linked to the operational definitions generated by the students during a class discussion. Additional research (by the students) was used to supplement the conceptual writing that the students are encouraged to perform. This writing helps to clarify ideas and construct knowledge according to Rivard (1994). The line of learning represents a planned instance that could allow metacognition about the process of writing:

> "The process of writing is important, not only for learning about something or acquiring knowledge, but for generating a personal response to something, for clarifying ideas, and for constructing knowledge." (Rivard, 1994, p. 970)

The journals provide opportunities for the students to reflect on their ability to learn science concepts and moreover, they allow the students to have a personal record of their own successes and failures with respect to the writing process. Reflecting on the strategies that they employ to write more cohesively and logically allows them to frame them more clearly in their minds for future instances and eventually becomes second nature. It is clear that time is required to allow these reflective moments to occur and that careful planning is needed during the implementation of the intervention process. It is envisaged that, with practice, students would perform metacognitive reflection outside class-times and in all facets of their education (not just in Science).

2.6 Writing Learning Strategies – Scaffolding/Modelling

In this research, a scaffold (in terms of writing learning strategies) represents a set of guidelines for students to follow in order for them to structure their writing in a coherent and logical progression. Modelling, in this sense, is the use of these guidelines to show the construction of a successful piece of writing. Warwick and Maloch (2003) suggest that a scaffolding approach is effective for improving scientific writing skills in the classroom but observe that this sometimes binds the students to a formulaic approach to writing. They identified the need for teachers to be 'reflective and responsive' when utilising the technique.

There is plenty of support for this technique. Porter, Guarienti, Brydon, Robb, Royston, Painter, Sutherland, Passmore and Smith (2010) suggest that low literacy skills may not be the problem with underachievement on writing tasks. It is possible that unfamiliarity with the requirements of scientific writing may be the issue. They provided guidelines (derived from Rutherford, 2007) to some of their students, as well as having a class discussion of the importance of each guideline and its relevance to scientific discourse. Rutherford's (2007, p. 11) guidelines included purpose and hypothesis, background and procedure (what do you already know), observations and data, errors and finally, organisation and conventions. Students who did not receive the guidelines often chose to represent the results of an experiment with a graph with little or no conclusive ideas whereas those who did often provided a more logical, evidence-based conclusion.

Rutherford herself (2007, p. 9) developed the guidelines due to her frustration with the "three-sentence conclusions, coupled with poor achievement scores". She found that, over the course of five formal sessions, where the guidelines had been given as a marking rubric, her students no longer needed the guidelines to write effectively by the end (pp. 12-13). Rutherford gave even more scaffolding to her special needs students with a “framed, or fill-in-the-blank, conclusion” (Rutherford, 2007, p. 13) with some success. This was based on work done by Terry (2001) for students with learning disabilities (Rutherford, 2007, p. 13).

The use of a guiding framework to illicit student responses is very similar to use of the Science Writing Heuristic (SWH) developed by Keys, Hand, Prain and Collins (1999). The SWH is a guide and scaffold for the writing process after experimentation in Science. It was used in a mixed methods study of middle-school students (Hand, Wallace & Yang, 2004). The specific activities used by the teacher during classes and practicals, combined with a writing template of scaffolding questions, saw a marked improvement in the complexity and quality of the students written responses. The scaffolding questions of the SWH are shown in Table 2.1 below. Hand, Wallace and Yang (2004) determined that student understanding of concepts improved along with a more coherent and developed writing style for students who used the Heuristic as opposed to those who didn’t.

Table 2.1 ***Scaffolding Questions of the SWH*** *(Adapted from Keys, Hand, Prain and Collins, 1999, p. 1069)*

Number	Title	Question
1	Beginning Ideas	What are my questions?
2	Tests	What did I do?
3	Observation	What did I see?
4	Claims	What can I claim?
5	Evidence	How do I know? Why am I making these claims?
6	Reading	How do my ideas compare with other ideas?
7	Reflection	How have my ideas changed?

LaConte and Berry (2006) suggested that the modelling of essential parts of a piece of scientific writing by the teacher is an essential part of improving students scientific writing. A science and literacy teacher respectively, LaConte and Berry (2006) collaborated in a study to improve students' writing in science. Many of the technical aspects of writing a persuasive text were covered by the literacy teacher with scientific examples provided by the science teacher. A discussion surrounding the necessary parts of a persuasive paragraph was designed to inform the students of good practice. After negotiation with the students to find a suitable topic, the teacher then "models writing an effective introductory sentence" (LaConte & Berry, 2006, p.n64) to start the students writing. The students then chose an argumentative side to the topic and T-charts were used to "organise the details on each side of the topic" (LaConte and Berry, 2006, p. 64). The teacher shows how to perform a 'think aloud' (verbalising the thought process for writing a persuasive sentence) before getting the students to try. This process focusses the students thinking behind the structure of the scientific writing.

Dweck (2000) identifies that students have two mindsets: growth (where students view their ability as something that can be changed) and fixed (where students view their ability as something that is static). In the face of challenges, students with growth mindsets relish them as learning opportunities regardless of the outcome, whereas students with fixed mindsets will see them as either 'easy' or 'impossible' tasks and react accordingly. She also makes links between these mindsets and effort. Students with an "entity theory-performance goal framework" (fixed mindset) believe that "if you have to work hard at something it means you're not good at it" (Dweck, 2000, p. 40). Conversely, students with an "incremental theory-learning goal framework" (growth mindset) believe that effort "allows you to fully use your ability and realize your potential" (Dweck, 2000, p. 40).

Scaffolding within a scientific writing task may help those students who are of fixed mindset and view the task as insurmountable. It separates the task into pieces which by themselves seem to require less effort than the whole. By breaking a task into manageable chunks, students may also find that they experience smaller, more-frequent successes:

> "Motivation based on goal setting involves a cognitive comparison process. By making self-satisfaction conditional on matching adopted goals, people give

direction to their behavior and create incentives to persist in their efforts until they fulfill their goals. They seek self-satisfaction from fulfilling valued goals and are prompted to intensify their efforts by discontent with substandard performances." (Bandura, 1993, p. 130)

Each part of the scaffold completed could represent a "cognized goal" (Bandura, 1993, p.n130) that increases the motivation of students who are unable to engage with the writing task due to its perceived enormity/complexity.

2.7 Summary

Chapter 2 was divided into five distinct sections. Section 2.1 introduced the literature review. Section 2.2 outlined the nature of action research as the paradigm used for this research project. Section 2.3 reviewed the idea of science writing as an area that many students find challenging, but as a potential gateway to identifying their understanding of scientific concepts. In Sections 2.4, 2.5 and 2.6, the strategies that can be used to improve students' writing and therefore their performance and self-efficacy were explored. Section 2.4 described the use of feedback in various studies of scientific writing and its role in informing students of what is being done well and where improvements could be made. In Section 2.5, examples of metacognition and its use as a way to encourage students to be reflective on their practice were described. Finally (in Section 2.6), the use of scaffolds and modelling to show students 'how to write successfully' in other research projects was outlined.

The next chapter describes the methodology that my research project utilised and the context of the research participants.

CHAPTER 3: METHODOLOGY

3.1 Introduction

In this chapter, the research methodology is discussed. Section 3.2 gives a brief overview of the research conducted in terms of the methods used. And Section 3.3 describes the context and nature of the population involved in the research. The next section (3.4) more deeply discusses the tools used to gather data for the research and the way in which the intervention strategies were implemented. The procedures used for data analysis are discussed in Section 3.5 and Section 3.6 describes the limitations of the project. The final section is devoted to a review of the ethical considerations required for the research.

3.2 Overview

In order to improve the students' writing in science, three main 'intervention strategies' were implemented in a year 6 class:

1) The use of a scaffold to guide their writing with analysis of strong and weak samples;
2) The use of peer/self-assessment for immediate feedback after the development of a co-constructed rubric;
3) The use of journals to document their progress and encourage metacognition of learning strategies.

In order to gauge the effectiveness of the intervention strategies, several students were interviewed before and after the implementation period. They were asked questions to gauge their attitudes and feelings about themselves as science students and as science writers. Work samples were collected from formative and summative tasks performed by the students early in the implementation period and compared with similar tasks performed later. The scientific writing tasks were designed to show students' ability to conclude from experimentation or second-hand information given. Most of this work was done in journals where the students could assess themselves and others. Metacognitive journal entries were also encouraged periodically although students could have written these at any time. The teacher/researcher also kept a journal over the course of the implementation period in order to reflect on the strategies being employed in the class and their effectiveness. These entries formed the basis of the reflexive research that documented the growth of the teacher, the observations of students' reactions to the strategies during their use and any issues surrounding the implementation of the strategies.

The data collection had three main time frames when information about the research could be obtained:

1) Prior to the implementation of the strategies
2) During the implementation of the strategies
3) After the implementation of the strategies.

Erickson (1998, p. 1159) suggests that qualitative data that is collected from many different sources and types is more effective and credible than data collected from one or just a few sources. For this reason, I decided to measure the effectiveness of the strategies by using four instruments (which are elaborated in sections 3.4.1 to 3.4.4):

a) Semi-structured interviews with students at different levels of achievement
b) Coded data/comments from journals used by the students in class.
c) Case study journal of the teacher.
d) Student scores in assessment items.

However not all of the instruments could be used in each time frame. Table 3.1 shows which instrument was used in each timeframe.

Overall the research took place over a 6-month period from February to July, 2013. Direct practice and use of the strategies averaged around half an hour every week for approximately 3 months in total. It should be noted that extra time was spent at the outset to set up the expectations and develop classroom procedures surrounding the strategies. Table 3.2 shows the relative timing of each phase of the project.

Table 3.1. *The Order of Use of Research Instruments*

Timeframe		Instrument
1	Prior	a) Interview, c) Teacher Journal, d) Scores on Work Samples
2	During	b) Student Journal c) Teacher Journal d) Scores on Work Samples
3	After	a) Interview, c) Teacher Journal, d) Scores on Work Samples

Table 3.2 *Relative Timing of Classroom Intervention Strategy Implementation and Measurements*

Order		Timeframe	Activity
1	Weeks 1-4	Prior	Interviews with selected students
2	Week 5	Prior	Diagnostic writing assessment to gauge initial levels
3	Week 6	During	Introduction to FRESH framework/scaffold (Strategy 1)
4	Week 7	During	Introduction to Journals (Strategy 3)
5	Week 8	During	Development of rubric and use for peer/self-assessment (Strategy 2)
6	Weeks 9-18	During	Use of strategies and informal diagnostic writing assessments to gauge interim levels (feedback into teaching)
7	Week 19	After	Diagnostic writing assessment to gauge final levels
8	Weeks 20-24	After	Interviews with selected students

3.3 Context of the Case

This study took place in a regional, coeducational, independent, K-12 school with about 1200 students. It is an advantaged school based on its rating of 1151 for the

Australian Curriculum Assessment and Reporting Authority (ACARA) MySchool ICSEA for which the average is about 1000 for other schools. The school profile on the ACARA MySchool website states:

> School Profile: (ACARA, 2015, accessed 4/2/2015)
>
> "At [the research study school], we value learning as the key attribute of developed individuals and communities. Students – their growth and achievement – are our work. As a school we believe that student wellbeing and self-efficacy are about learning and progress. Each student must be deeply engaged in their learning and understand that how they think and behave affects their potential for optimum development. We work to support students to determine ambitious and achievable goals and to use challenges and set-backs as opportunities for learning. We work to support students to believe that, with proper instruction, support and coaching and with persistence and effort, they can make optimum progress in whatever they choose to do in life. At [the research study school], we believe that students can only make optimum progress in an environment in which they are encouraged and supported by those around them.
>
> In order for students to make optimum progress, the most important resource is the quality of teaching. It is our responsibility to ensure that every opportunity for learning, in every context, for every student, is maximised, directed and supported. [the research study school] is committed to continuous improvement in teaching practice. To deliver on this commitment, significant resources are allocated to both maintaining an exemplary standard of practice and to the identification and implementation of evidenced-based teaching approaches proven to be the most effective in improving student learning outcomes…..
>
> …..The curriculum provides students with diverse opportunities: from physical fitness and team sports, musical ensembles and theatrical productions to international study tours to Boston, San Francisco and Cambridge and involvement in local and overseas service activities, particularly our humanitarian work in Papua New Guinea.

> [The research study school] is a co-educational day and boarding school, enrolling students from Early Learning to Year 12."

During the year in which the research was conducted, the research school was performing similarly or above 'like' schools in all areas of the National Assessment Program – Literacy and Numeracy (NAPLAN). The school's NAPLAN data on the MySchool website (ACARA, 2015, accessed 4/2/2015) is summarised in Table 3.3. The descriptors "above" and "substantially above" are based on the comparisons made by ACARA. The ranges used for these descriptors are not provided on the ACARA website.

Research was conducted in a Year 6 class comprising of 19 students (10 females, 9 males). They were selected due to their relative unfamiliarity with science and science writing. They had previously had one year of science taught by dedicated science teachers. The entire class participated in the research, although only eight students (4 female, 4 male) were selected to be interviewed. These eight students were selected based on their gender (equal numbers of males and females) and their relative ability level based on previous grades. Care was taken to select students from the top, middle and bottom bands of ability. This was to allow comparison by gender and by entry 'rank' during the data analysis.

The students shared a familiarity with each other due to the fact of being in many of the same classes together and sharing the same home-room where they would meet at least once a day. Most of the students had known each other for at least one year prior to the research project. This is an important point as their comfort with each other meant that they felt more likely to share their thoughts in peer assessment and contribute to class discussions. It also meant that many friendship groups had been established in the class, which could influence the nature of their feedback to one another.

Table 3.3 ***Summary of NAPLAN Results of the Research School Compared to Australian Average Achievement***

Year	Reading	Persuasive Writing	Spelling	Grammar and punctuation	Numeracy
3	Substantially above	Substantially above	Substantially above	Substantially above	Substantially above
5	Substantially above	Substantially above	Substantially above	Substantially above	Substantially above
7	Substantially above	Substantially above	Above	Substantially above	Substantially above
9	Substantially above	Substantially above	Above	Substantially above	Substantially above

The researcher was also the teacher for the class and had taught Science for ten years prior to the research project. The male teacher/researcher had completed a Graduate Diploma of Education and a Bachelor of Science in Geophysics. The classroom dynamics could be described as congenial, with very few issues due to behavioural problems or personality conflicts. Prior to the start of the research, the researcher had taught the participating students for approximately a fortnight and expectations around classroom procedures and etiquette were still being developed.

Table 3.4 displays the names (pseudonyms), gender and relative placement of the students academically based on their percentage scores in common assessment tasks. The table has been sorted by the average of four assessments the students had performed before the research began.

Table 3.4 ***Pseudonyms, Gender and Academic Ranking of the Research Group***

Pseudonym	Gender	Biology 1	Graphing	Biology 2	Chemistry	Average
Ralph	Male	41	27	48	20	34
Matt	Male	23	40	44	52	40
David	Male	64	40	48	44	49
Michelle	Female	68	40	60	28	49
Caleb	Male	68	47	40	60	54
Eleanor	Female	73	67	48	28	54
Karen	Female	73	53	48	44	55
Badley	Male	73	40	64	48	56
Steven	Male	64	60	60	68	63
Peter	Male	86	60	68	72	72
Beth	Female	82	60	88	60	73
Allison	Female	82	73	68	80	76
Rowena	Female	73	73	80	88	79
Nathan	Male	91	67	76	84	80
Lucy	Female	91	73	84	76	81
Abigail	Female	86	73	88	84	83
Arthur	Male	91	80	84	80	84
Jenny	Female	77	93	92	76	85
Tina	Female	95	80	88	Absent	88

3.4 Method and Instruments

Polkinghorne (2007, p. 475-476, quoting Isaac & Michael, 1987 and Campbell & Stanley, 1963) summarises the use of instruments in social science research and their validity in the quote below:

> "Conventional social science research makes claims that an instrument measures variations among participants in a theoretical construct. It attempts to produce convincing or validating evidence that the instrument's content accurately samples the kinds of things that make up the construct (the content validity); that its measurements are consistent with other instruments which measure the same construct (criterion-related validity), and that the instrument is actually measuring the concepts and qualities of the construct (construct validity; Isaac & Michael, 1987). It also makes claims that a particular intervention is the cause of measured results. The claimed results

> are that the mean scores from sample groups of people most likely did not occur by chance. To justify these kinds of claims, it attempts to eliminate all possible explanations other than the intervention that could account for the measured results (internal validity; Campbell & Stanley, 1963). It also makes claims that its findings apply not only to the participants who were examined but also to all those in a comparable population. It uses randomized selection of participants and statistical procedures to convince readers of the validity of these generalized claims."

In terms of the content validity, the researcher believes that the improvement in scientific writing of the students after the intervention (the construct) was adequately measured by the four instruments described below. As each instrument was found to produce similar indications of improvement, they satisfy criterion-related validity. The internal validity is explored in the Chapter 4 (Findings) and the construct validity is analysed in the final chapter (Conclusion and Implications).

Each instrument is elaborated below.

3.4.1 Interviews

Eight students of mixed gender were interviewed from the Year 6 class. They had been identified as strong, medium and weak science students relative to their peers based on previous assessment and anecdotal records from the class teacher. Students were asked questions regarding their performance as science students, as science writers and as scientists. Following the use of the strategies in class a follow-up interview with those students gauged any improvement in their views of themselves and self-efficacy.

Each interview was a one-on-one semi–structured interview which allowed the personal feelings, beliefs and attitudes held by the participant to be captured at a point in time. The open-ended nature of the information gathered was considered to be the main advantage of this instrument. It allowed a range of answers and it often included follow-up questions to clarify/expand particular areas of interest.

The quality or usefulness of data obtained from the interview required some insight and quick wittedness on the behalf of the researcher to follow interesting threads, gauge participant comfort levels and maintain impartiality:

> "The interviewer has to simultaneously reflect on the information being provided in order to relate it to prior information; plan the next question; decide whether to pose the question or make provision for the respondent to answer it in his or her own time. Most importantly the respondent has to be convinced that the interviewer is not an adversary but is at the very least impartial…Failure to attend to these basic principles of interviewing will diminish the quantity of data obtained" (Partington, 2001, p. 43).

Anderson, Nashon and Thomas (2007, pp. 181-182) note that metacognition has no central definition within the science education literature but a common thread is that it involves the learner being aware of how he/she learns. The post-intervention interviews had elements of 'metacognition' in this sense as the students were asked to evaluate which strategy had helped them the most to (hopefully) improve their science writing.

The questions for the semi-structured interviews were based on the prompts in Table 3.4 below. They were divided into two distinct sections. The questions in section 1 were chosen to elicit response about the students' attitudes towards science and their general self-efficacy. The questions in section 2 were chosen to get the students' ideas about the writing they had done in science before and after the intervention strategies were put in place.

Table 3.5 ***Interview Prompts for the Semi-structured Interviews***

Section		Prompts
PRE-INTERVIEW	Section 1 View of self as a science learner	➢ What is ‘science’ to you? ➢ How do you view yourself as a scientist or science student? ➢ Do you find science difficult? Can you describe why/ why not? ➢ What do you find most enjoyable about science? ➢ What do you find the least enjoyable about science?
	Section 2 View of self as a science writer	➢ In science, we often do an experiment and then try to communicate what we have found out in a scientific report. How challenging do you find this writing? ➢ Can you describe why/ why not? ➢ Do you think your writing is of a high standard? ➢ Can you describe why/ why not (try not to think in terms of the marks you get, but the actual content)?
POST-INTERVIEW	Section 1 View of self as a science learner	➢ Do you feel as though your perception of yourself as a scientist has changed at all this term? ➢ Can you describe why/why not? ➢ What do you find easier to do in science after a term of work? ➢ What do you still find challenging?
	Section 2 View of self as a science writer	We have been using different methods in class to try and improve how we think about writing about experiments and scientific concepts. They were: 1) Scaffolded framework for writing assessments in the form of a series of ‘success criteria’. It was

		visible during practice tasks and for modelling examples. 2) Peer/self-assessment of practice tasks. 3) Journaling to reflect on experiences with writing assessments, identifying what was productive and what strategies help. ➢ What key things stick in your mind about the methods? ➢ Have they helped you at all? ➢ Did you find one more useful than the others? Why? ➢ Do you think your writing is of a high standard? ➢ Can you describe why/ why not (try not to think in terms of the marks you get, but the actual content)?

3.4.2 Student Journal

The 19 students in the research group were issued with a journal in which they could practise writing conclusions to experiments, peer/self-assess their own work, and be encouraged to write about the experience of improvement (metacognition). Glogger et al. (2012, p. 452) say that learning journals can 'deepen understanding and retention' and 'apply cognitive and metacognitive learning strategies' through their use as a part of coursework. As well as conclusion writing samples, direct quotations were taken from the journals as samples of their metacognitive thought processes during the implementation of the intervention strategies as well as their peer or self-assessment feedback.

As part of the peer/self-assessment, the students and teacher/researcher co-constructed a 'marking' rubric to help the students to measure their progress. It was created using a simple grid consisting of the five criteria of the writing acronym/scaffold 'F.R.E.S.H.' (elaborated on in the next sections) and whether expectations were, below, at and above the required level/standard deemed adequate by the class. Samples of these were collected with the journals. The co-constructed rubric for 'conclusion writing' is given in Table 3.5.

3.4.3 Teacher Journal

The teacher journal was a series of reflections completed by the teacher/researcher during the implementation of the strategies in an attempt to improve the students writing. This is mainly to enable answering of the 4th research question (What problems arise from the implementation of the intervention strategies?) in detail rather than infer it from the other data sources. It is comprised of the lesson plans and subsequent reflections of the lessons corresponding to the research based on informal observations made of the class during its implementation. It is also part of the 'feedback loop' that informed the teaching and guided the use and order of how the strategies were 'unfurled' in the classroom. Arends (2001, p. 26) suggests that a reflective teacher is an effective teacher and "learns to approach unique situations with a problem-solving orientation and learn the art of teaching through reflection on their own practise". As the researcher/teacher was an active part of the project, the reflections provided an insight into the evolutionary nature of the research, and how the thoughts of the researcher changed over time.

Table 3.6: ***Co-constructed Rubric for Writing a Conclusion by the Researcher and the Research Class***

Not at required level	At level	Above required level
Hypothesis evaluation not included	Hypothesis is evaluated as right/wrong	A general statement is used to describe the outcome of the experiment
Evidence is not used to support the findings	Observations or numerical data are used to support the findings	Multiple sets of data or observations are used to support the findings and demonstrate any relationships
Scientific theory is not mentioned	A brief explanation of reasons for the data is given	An in depth discussion of scientific theory is used to describe why particular data was obtained
Reliability is not mentioned	The reliability of the data/observations is discussed	The reliability of the data/observations is described in detail and as well as the effect of error on the conclusion’s validity
Further experimentation is not mentioned	Suggestions that may further the understanding of the aim are included	Modifications to the method to improve validity are suggested as well as other variables to change

3.4.4 Scores on Work Samples

Numerical scores were obtained to ascertain the level of achievement for scientific writing (conclusion) for samples of work taken before, during and after the implementation of the strategies aimed at improving the students' writing. These scores would eventually become part of the students' overall grade once amalgamated with other testing. For the purposes of the research, the final grades that the students received were not fine-grained enough to be truly insightful and therefore the raw scores for the scientific writing in the journals were considered only. The samples of work considered were taken from the four tasks given in Appendix C. They are given in order of their use, with Task 1 conducted before the intervention process began and Task 2 conducted shortly after the introduction of the 'F.R.E.S.H.' scaffold. Task 3 was conducted after some Self and Peer Assessment practise, and Task 4 was conducted at the end of the intervention phase.

The assessments of the writing samples in the journals were considered to be embedded assessments (Treagust, 2001, quoting Gallagher, Parker & Ngwenya, 1999 and Wiggins, 1998, p. 138) as the skill was assessed in order to inform the teaching in a cycle. This 'feedback loop' allowed the intervention strategies to be modified or reworked to cater for the needs of individual students or groups.

The level of achievement was based on an outcome statement that the research school had adapted from many sources. The 'Conclusion Continuum' is presented in Table 3.6. Each level of achievement is accompanied by a description of the criteria that are required to be awarded that level. The criteria were developed by Science teachers at the research school to help assess skills such as writing conclusions, drawing graphs, etc. The F.R.E.S.H. scaffold was developed to help students achieve higher on this assessment scheme by addressing some of the central ideas that weave their way through the criteria on the Conclusion Continuum.

Table 3.7. ***2013 Conclusion Writing Continuum***

Level	Level descriptor
9	The student is able to produce a clear, well supported conclusion which is conveyed in a highly persuasive and appropriate manner. The student develops a scientific principle that is accurate and clear and evaluates the aim and hypothesis and demonstrates an **advanced conceptual understanding**. The student evaluates the data in terms of **its reliability, citing specific examples**, and the effect this has on the generalisations drawn from the data. The student is able to recommend appropriate tests that should be carried out next to test their new hypotheses which will **further their understandings that are associated with the aim**, **and improve the reliability.**
8	The student is able to produce a clear conclusion which is conveyed in a highly persuasive and appropriate manner. The student develops a **more advanced scientific principle** that is accurate and clear and evaluates the aim and hypothesis. The student is able to evaluate the data in terms of its **reliability, citing specific examples**, and the effect this has on generalisations drawn from the data. The student is able to make broad statements about **further tests that should be carried out** next to test their new hypotheses.
7	The student is able to produce a clear conclusion which is accurate and appropriate with an accurate evaluation of the strength of the evidence. The student is able to develop a **basic scientific principle** from the results collected and the hypothesis and aim are evaluated accurately. The student is able to evaluate the data in terms of its **reliability in broad terms** and the effect this has on generalisations drawn from the data. The student is able to make broad statements about **further tests that should be carried out** next to test their new hypotheses.

6	The student is able to produce a clear conclusion which is accurate and appropriate with an accurate evaluation of the strength of the evidence. The student is able to develop a **basic scientific principl**e and the hypothesis and aim are evaluated accurately.
5	The student is able produce a clear conclusion which contains an accurate summative statement and a relatively accurate evaluation of the strength of the evidence. The student attempts to give a **general scientific principle** based on the results, but it is **not clear** or is not entirely correct even based on the experimental information.
4	The student is able to produce a clear conclusion which is relevant to most pieces of evidence. The student includes accurate summative statements and attempts to show this **using evidence or data** from the experiment to exemplify the statement.
3	The student is able to produce a conclusion which consists of a relatively short summary of the argument and a basic summative statement. The student **addresses the contention, aim and hypothesis** in the conclusion. The student is unable to support the summative statement with evidence
2	The student is able to produce a conclusion which consists of a simple restatement of the contention or aim. Although the student recognises the need for addressing both the hypothesis and providing a summative statement, the **conclusion rarely includes both**.
1	The student is able to produce a conclusion which consists of a simple **restatement of the contention or aim**, but is unable to assess the hypothesis or the strength of the evidence

The three main intervention strategies used within the Year 6 Classroom for the research are now elaborated.

The first intervention strategy was the use of a scaffolded framework for scientific writing that was introduced and elaborated with the students in the form of a series of ‘success criteria’ similar to those described by Porter et al. (2010).

Over the course of several years, the Science teachers at the research school had already developed several acronyms to help students remember to the success criteria required for a specific piece of scientific writing. These memory devices allowed students to remember the criteria for the creation of a hypothesis, the writing of a reliable method, the analysis of data and (most importantly for this research) the writing of a scientific conclusion. These were created in small teams of collaborating teachers and shared with the entire department.

The conclusion scaffold became known as ‘F.R.E.S.H.’ to staff and students alike. The meaning of each letter is described in Table 3.7. Students were asked to remember that the letters of the acronym were not in the order that a conclusion should be written. The mnemonic was only given to help them remember what needed to be included. If it was presented in writing order it should be H. E. S. R. F., hence it was almost ‘reversed’ from F.R.E.S.H.

Table 3.8 ***The F.R.E.S.H. Scaffold/Acronym***

Letter	Meaning
F	Further Experimentation
R	Reliability of Data
E	Evidence
S	Scientific Theory
H	Hypothesis Correct?

It bears striking similarities to the SWH developed by Keys, Hand, Prain and Collins (1999, p. 1069) in terms of how the tested hypothesis (claim) is supported with evidence from the experiment and compared to other ideas (scientific theory). This is coincidence as the F.R.E.S.H. scaffold was developed independently by the Science teachers at the research school.

The framework was visible during class time and during assessments in the form of a large poster in plain sight. The framework was to be referred to during formative tasks and for modelling of exemplars with the students as per LaConte and Berry (2006). The poster is recreated in Figure 3.8 and outlines the significance of each letter in the acronym.

Figure 3.1 ***The F.R.E.S.H. Poster***

Further Experimentation

- What would improve the accuracy of the experiment?
- What would additional testing would further our understanding of the aim?

Reliability of data

- How reliable is the data in the experiment?
- Were there outliers?
- Does the reliability affect the overall accuracy of the conclusion?

Evidence

- What are some actual pieces of data that support the generalisation made from the experiment?

Science Theory

- Why did the results happen the way they did?
- What scientific concepts explain these results?

Hypothesis and generalisation

- Was the hypothesis correct/incorrect/inconclusive?
- What is the overall relationship that the results imply?

The scaffold was presented to the students in the very first class after an initial attempt at writing a conclusion and it became a central theme of the entire research project. It was referred to in order to create a peer assessment rubric, to write strong and weak samples and to consider as a writing strategy during reflections.

During one of the initial classes, strong samples based on the criteria from Table 3.6 were co-constructed with the students utilising the F.R.E.S.H. scaffold and reflected upon. An example of a strong work sample completed with the class is given in Appendix B. This was considered to be similar to the modelled writing described by LaConte and Berry (2006). Another modelled writing task involved the provision of strong and weak examples of writing for each part of the acronym so that students could select which piece of writing was the best one (see appendix B). This is in line with the 'flawed samples' work done by Brandt (1971). Students also practised the use of the rubric outlined in the second strategy.

The second strategy was the use of peer/self-assessment with respect to the framework developed with the students in strategy one. Some class-time was to be spent developing guidelines and a rubric for students to follow, similar to the research performed by Butler and Nesbit (2008). The rubric was shown above in Table 3.5. After conclusion writing samples were written, the students were asked to peer/self-assess their work using this rubric. Sometimes opportunities arose for them to 'retry' the conclusion based on the feedback they received.

The third strategy employed was the use of metacognitive writing opportunities in the students' journal (similar to that used by Conners, 2007) The students were encouraged to reflect on their experiences with scientific writing, hopefully identifying what was productive and what strategies helped. The journal was introduced as such to the class, and students were encouraged to put all of their comments into the journal. A couple of 'formal times' (time specifically dedicated to metacognition) were given halfway and at the end of the implementation process for students to share how they thought their progress was going and which strategies worked for them.

3.5 Proposed Analysis of Results

Mason (1994, p. 91) suggests that qualitative data such as interview transcripts can be analysed using different techniques. One of these was “to search the data set for themes, to develop analytical categories, and to index the data accordingly.” But she also warns that “deciding when someone is talking about a particular topic is often a matter of interpretation”. It is with this in mind that the data collected from the three qualitative instruments (the interviews, the student journals and the researcher journal) was analysed thematically, with commonalities between student responses used to support the assumptions made about the effectiveness of the intervention strategies. The interviews were considered to be narratives of the students’ development over time given their chronological difference (before and after the interventions). Bailey and Jackson (2003, p.n59, quoting Polkinghorne, 1995 and Glaser and Strauss, 1967) summarise the analysis of narratives in the following quote:

> “Analysis of narratives is produced using paradigmatic reasoning, a logical way of knowing that uses classification systems to bring order to various elements in the world. Paradigmatic reasoning refers to the way people understand the world through “cognitive networks of concepts” (Polkinghorne, 1995, p. 10). Analysis of narratives which emerges from paradigmatic reasoning reduces stories to their common elements thus producing general knowledge. Using this cognitive mode of thought, the researcher scrutinizes the data to discover categories describing common themes that appear across the stories and then through further analysis draws relationships between the categories. Polkinghorne points out that analysis of narratives requires methods similar to those proposed by Glaser and Strauss (1967) in their grounded theory method.”

Although the grounded theory method Glaser and Strauss (1967) has obvious relevance to the data analysis of the interviews performed in this thesis, due to the small sample size, complicated coding of the data was not considered necessary by the researcher to extract themes common to the students’ interview transcripts. Instead, themes were identified by considering the responses to the stimuli in the semi-structured interview in turn. It was readily apparent that the students often held similar views and their responses were grouped together under categories such as "enjoyment linked to science vocabulary". The categories are further explained in the data analysis chapter.

Similarly, the student journals were read for commonality in terms of themes shared by the students. Supporting quotes were transcribed from the journals and placed into categories. Again, due to the small sample size, it was not necessary to code the comments to draw out the most common themes.

The teacher/researcher journal was used primarily as a single case-study and relevant quotes were drawn to emphasise particular points that were relevant to the research questions, or to support the themes identified from the student interviews and journal entries.

Quantitative data was generated by selecting four common pieces of scientific writing in the form of conclusions that the majority of the students in the class had completed. They were assessed against the numerical levels of achievement as given in Table 3.5. This data was put into Microsoft Excel so that numerical manipulation could be performed. The mean of the students' performance was calculated by task in order to identify any improvement in the standard of their writing over the course of the intervention. Additional sorting permitted consideration of differences by gender, and by overall ability levels based on their starting point.

3.6 Some Limitations of the Research Design

Open-ended assessment tasks such as writing a conclusion for an experiment can limit students' achievement if they are poorly designed. Tasks which have inaccessible areas and aren't open 'enough' will not be useful for gauging a students' true potential. This means that as formative assessment, it will lose value as the goal is to engage with students' zone of proximal development which is "the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (Vygostky, 1978, p. 86). Some critical analysis of the conclusion writing tasks was necessary to ensure that they were open enough to allow entry by the most capable and least capable writers in the class.

It has been shown that there can also be reluctance to change practice to an open and transparent assessment for which students are aware of what is required for achievement, because the teacher wanting to remain 'secretive' about the criteria (Simon, 1992, p. 361). The co-constructed rubric for gauging the effectiveness of the conclusion writing meant that the marking criteria were open and transparent by their very intent. The presence of the F.R.E.S.H. acronym as scaffold/success criteria in the room meant that the students knew exactly what was being asked for during all of the writing samples used in the quantitative research instrument.

Why use a semi-structured interview? A semi-structured interview is "based on the use of an interview guide…a written list of questions and topics that need to be covered in a particular order" (Russell, 2011, pp. 157-158). It allows for more open-ended responses while still directing the respondents to address common themes. Its limitations are that some interesting conversations that are off topic may never be held, and that it becomes more difficult to code the responses in a quantitative way. This is most suited to the case-study format given that the relatively small sample size renders a quantitative coding method unreliable. Also, given the age of the respondents, the prompting questions were probably necessary to elicit more mature and reflective responses. Care must be taken, as Partington (2001, p. 34) says, to ensure that the cognitive and social framework is not too sophisticated or "a child may feign understanding and provide answers that seem to satisfy the interviewer".

Russell (2011, pp. 158-165) gives several other important guidelines on interview techniques such as fostering respondent motivation, probing for more information and focussing respondents who say too much or not enough. Partington (2001, p. 35) suggests that empathy and rapport are essential to get the most useable information from the respondents. It is assumed that, as the respondents were in a student-teacher relationship already, there would already be established rapport with them. However, care was taken to ensure that the students did not feel pressured to answer in a particular way.

The following statements were used as a pre-interview discussion/preamble:

> "I am researching how to improve students' writing in science. I am interested in your point of view, because your input will help me to understand how effective some teaching strategies have been. I want to know about two main areas – how you view yourself as a learner and how learning strategies may have helped you. I want to use your responses because YOUR opinions and observations are important to my research. I would like to record your responses but assure you that you will be completely anonymous and that this will not affect your grades nor anything else in class."

According to Denzin and Lincoln (1994, p. 372) a danger associated with less structured interviews is the tendency for the interviewer to pretend to be an 'invisible participant' when the reality is that the interviewer does influence the quantity, quality and accuracy of the responses by his/her interactions with the respondent. Body language, use of technical terms, facial expression, listening to responses (to guide the next question) are just some of the ways in which the interviewer can influence responses. It was therefore important to be reflexive when interpreting the interview transcript. Some reflection was needed after each interview to enable this reflexivity.

3.7 Ethical Considerations

Ethics approval was granted by the Curtin University Human Research Ethics Committee (approval number: SMEC-39-12) before the project commenced at the research school and data was collected.

Ethics is the set of guidelines on how you should interact with the world so as to ensure that no-one is harmed or disadvantaged by your actions. In educational research, it is the researcher(s) who must adhere to these guidelines as set by the rules and expectations of governing bodies such as educational institutions, regulatory bodies and the government. History has shown that research without ethical considerations can be harmful to the participants both emotionally and physically (Anderson, 1998, p. 16 &17).

Ethical educational research should include the basic tenets of integrity, respect, beneficence and justice. The researcher should be unbiased and respectful of the participants and concerned parties in the planning, conducting and presentation of the research. The participants should not be made to feel emotionally or be physically harmed during the course of the research. The research should not burden the participants without due reimbursement of some form.

Einarsdóttir (2007, p. 204) describes how the relationship formed between researcher and participant is often fraught with ethical issues, especially in more qualitative methodologies. She shows how power inequities, perceived consent, and accurate transmission of the participants' views or intent are just a few examples of the numerous intricacies that the researcher must reflect upon and consider through each phase of his or her ethically-bound research.

The sample included students who already have a 'student-teacher' relationship with the researcher. The nature of the instruments required interaction with some participant on a personal level in a close one-on-one interview as well as at 'a distance' (the journals and work samples). These interactions required a framework of ethical guidelines to overarch the research process from its outset to its conclusion.

The participants were required to read an information sheet about the research and sign a consent form that says that they understand what the research is for and their role in the research (See Appendix A). Power relationships such as that from a teacher to student or from manager to teacher could influence the decision to participate in a study (Einarsdóttir, 2007, p. 204). Because the researchers' request for participation was seen as coming from an authoritative figure who had influence over grades and reporting, this may have changed students' decisions as to whether to participate or not. It was made clear to the students involved that participation was entirely voluntary. The aim and methods of investigation were made clear to the participants and their parent/guardians so that they knew what they were signing up for and what they should expect in return. With this in mind, the researcher was open with the hypothesis that the students had room for improvement in science writing and that the intervention strategies would benefit them. Being open about this preconceived idea was necessary to maintain integrity in the research.

Consent was also needed from the management of the school to use the facilities and students for the research. The expected improved outcomes were communicated and deemed well worth the extra time spent in class with the students.

As previously stated, the data collected was mostly qualitative in nature. Some of the statements collected were bound to be emotive in nature and require sensitive handling. The participants understood that their responses were completely anonymous (or involved using a pseudonym to preserve narrative style) and would in no way be linked to them. They also had the power of veto over what may or may not be used in the research. The participants were made to feel that they had total control over their voice in the research. The responses to interview questions were analysed to ensure that there was no information that can be used to identify the owner such as historically-significant data or character-specific references (Clark & Scharf, 2007).

Considerable planning ensured that the questions in the interview were phrased in such a way as to probe, but not make the subject feel threatened, undermined or attacked. Care was taken to be considerate of ethnicity, gender, experience and other background characteristics that could have potential to offend or engineer emotional harm. Clark and Scharf (2007, p. 402) point out that interpretive methodologies that gather personal data such as interviews and case studies have potential for causing harm due to "the consequences of truth" (p. 399). It was possible, in the interview, that the researcher might have uncovered feelings or experiences that possibly compromise the use of that data in the research as it can change the relationship between researcher and participants by going beyond the realms of the consent given prior to the interview due to the 'seduction of the caring interview', to paraphrase Clark and Scharf (2007, p 405). Consequently the use of this qualitative data was subject to review by both the researcher and participant before it could be included in the research.

Etherington (2007, p. 601) illustrates how consent is not a 'one-off' thing at the beginning of the research process. She describes an interview in which, as potentially harmful emotions and thought processes are experienced by the participants, she offers the participants the opportunity to withdraw from the process (pp. 603-611). As part of the consent agreement before research begins, the researcher made make it clear that the participants may withdraw from the process completely and have ‘ownership’ of their comments. The researcher was reflexive during the interviews and mindful of comments that may have required a reminder that withdrawal was available. It required the researcher to be quick thinking but also mindful of the potential for harm that an unplanned divergence from the topic may have led to. The timing of the interviews neither affected the students’ normal schooling time nor adversely affected the school’s classroom or office usage.

The interviews were recorded and transcribed by the researcher. Because often the ‘way’ something is said can have different meaning when converted to text, care was taken to ensure that the comments included (for an audit trail) were presented in the way the participant intended. Descriptions of emphasis, tone and speed are ways that the text were given more emotive context when required. True interpretations of participant mood and intent were acquired by asking the participants to feedback on the transcriptions’ representation of their words (or lack thereof). The participants were allowed to see and comment on all drafts of the thesis, with continued right of informed consent. They were able to make their choice of withdrawal after reviewing how their comments were interpreted. They were able to assess whether there would be any social or emotional harm to them upon submission of the thesis.

As the researcher was a member of the teaching staff at the school, the research had to take place in a way that did not detract from the regular curriculum taught and did not adversely affect the teaching quality and responsibilities of the researcher and learning experiences of the participating students. Communication between the researcher and the school was maintained throughout the research so as to be completely transparent during the interviews and implementation of the improvement strategies. The school had the right to withdraw its support for the research should it feel the need. The strategies were made available to other classes through common/shared lesson plans so as not to disadvantage other students in the year level cohort.

Acknowledgement of the students involved (anonymously) and their educational organisation within the text of the research is another way of repaying the debt owed them for their participation. The school was given the choice as to whether they wish to be acknowledged as having taken part in the research in the thesis or any publications that follow. The school was also given the right to inspect and withhold any data collected from their students and consequent analysis. It was expected that, in return for its participation, the school would receive a copy of the report to use as it sees fit. It is also expected that the students who participated in the research would benefit from the extra attention given to them (over and above the normal amount given to them by the classroom teacher).

The researcher initially required access to prior results of students at the researcher's school. Because an interview area convenient for willing participants and the researcher and the school was needed, a school meeting room was utilised after-hours for this purpose. The interviews were recorded onto a computer and an external microphone. The electronic data collected during the study was stored on a computer protected by passwords. Any paper format data collected was stored in a locked filing cabinet. All electronic and paper format data produced were be stored in a safe and secure location in the Science and Mathematics Education Centre at Curtin University for a period of 5 years after the publication of this thesis.

3.8 Summary

Chapter 3 provided the reader with the methods and instruments that were utilised in the research. Three major, but interwoven, strategies were employed to improve the writing of a class of Year 6 Science students with a focus on conclusions drawn from experimental data. The success of these strategies was measured by interviewing selected students, analysing writing samples and reviewing student journals and the journal kept by the teacher. The data analysis procedures were defined and the limitations and ethical considerations were also described in this chapter.

Chapter 4 provides the results and analysis of the data collected. It is divided into distinct sections which consider the data collected from each instrument separately.

CHAPTER 4: FINDINGS

4.1 Introduction

This chapter is divided into four sections that show the findings from each instrument and their subsequent interpretation. Section 4.2 addresses the data obtained from the interviews and consists of quotations from the students before and after the implementation of the intervention strategies. Section 4.3 describes the information contained within the learning journals that the entire class maintained during the project. In Section 4.4 the researchers journal is analysed from the perspectives of planning, conducting and reflecting on the strategies. Section 4.5 examines the work samples and the quantitative data obtained by assessing the level of ability of those samples. Section 4.6 provides a summary of the chapter.

4.2 The interviews

The semi-structured interview were conducted in two parts, one before the implementation of the intervention strategies and one after. In all of the interview quotations, the names of the student have been changed for privacy. The students' 'cases' are presented with some background information deemed relative to each student as he/she appears in the discussion. As described in Section 3.5, the transcripts of each interview were analysed for elements of commonality which are presented as 'themes'. Each theme is presented with supporting quotes from the interviews. The style of the thesis becomes slightly narrative in nature as the journey of the students through the intervention process is presented in the next few sub-sections.

4.2.1 Theme 1 – Quantity Versus Quality

Bradley was a quiet student who was often distracted in class. His grades in science to this point had been fairly low and he found writing to be a chore. In the initial interviews, one of the themes that emerged was that students seemed to link quantity to quality. The initial interview with Bradley was quite challenging for the researcher as many of his answers were one word replies and he required some 'prodding' to elaborate on his answers. However, Bradley's interview showed a theme that was consistent with other students - their impressions of writing in science contexts was that the more information that was contained in the writing, the better they perceived it to be:

> *RESEARCHER: That's okay, maybe we should change this. Do you find science interesting?*
> *BRADLEY: Yeah….*
> *RESEARCHER: Do you find it challenging at all?*
> *BRADLEY: Sometimes…..*
> *RESEARCHER: What do you find challenging?*
> *BRADLEY: Like, writing reflections sometimes, like really long ones.*
> *RESEARCHER: So it's the length of the writing that's hard to do?*
> *BRADLEY: Yeah.*
> *RESEARCHER: What is it about the 'length' do you think? Are you just not sure what to write? Or*
> *BRADLEY: Well like, I write everything that I know, and then like, other people have way longer….*

Bradley was very aware of his position in the class in terms of his understanding of science and his rank compared to other students. He often sat next to students who were relatively talented in Science and he was able to see that those students often wrote a lot more than he did on most tasks. It is readily apparent that 'length' and 'good' formed a link in Bradley's mind.

Very similar to Bradley, David found science writing to be quite challenging. David tended to be more vocal than Bradley in the interviews, although in class he was often quiet and found it hard to contribute to class discussions without direct questioning. David made the link between quantity and quality as well:

> *RESEARCHER: That's fine… do you think your writing is of a high standard?*
> *DAVID: Not a high standard, I would probably….like, sometimes it's a low standard or a medium standard because I'm not putting in a lot of effort or all the information in.*
> *RESEARCHER: So it's about information?*

DAVID: Yeah. It is….

Interestingly, 'information' seemed to be a general term to describe data, theory, vocabulary and anything related to the subject being written about. Some of the students were a little more articulate about the nature of 'information' but still made the 'more is better' link. In the next case, another boy named Arthur makes a specific link to data as 'statistics'. Arthur is one of the top students in the class and has consistently achieved high grades. He is a very fluent writer but requires some guidance regarding structure:

RESEARCHER:Do you think your writing is of a high standard?
ARTHUR: Ahh, yes.
RESEARCHER: What makes you think that?
ARTHUR: Because I include lots of statistics and basically everything

It was not uncommon for Arthur to submit half a page of writing, whereas David and Bradley might only turn in a sentence or two. Although Arthur was submitting a lot of work, it may not necessarily have been succinct nor have all of the required parts to make it a logical piece.

Eleanor was a girl who struggled to master the scientific concepts in class, and was often distracted by her friends. Eleanor was in the same situation as Bradley and David, finding it hard to write after performing an experiment. She had a low opinion of herself in terms of her conclusion writing. A sense of frustration was evident in her interview:

RESEARCHER: How good do you think you are at explaining why something has happened?
ELEANOR: Um…I wouldn't say really good
RESEARCHER: Why not? What holds you back do you think?
ELEANOR: Well you know what the experiment has done. But you just go back and you don't write every single little thing.

Eleanor made an important point in that she often would recognise the outcome of the experiment but had trouble putting it down on paper, but she had not made the connection between being able to say what has happened and the reason why it had happened.

4.2.2 Theme 2 - Vocabulary Issues

As mentioned in Chapter 2, links have been made between mastering the vocabulary of science and coherent writing. This theme was evident in the interviews with the students both after and before the intervention strategies. In the initial interview with Eleanor, she seemed hesitant to make Science seem like anything but her favourite subject. Upon a little more questioning, she opened up a little:

> *RESEARCHER: What do you find enjoyable about science?*
> *ELEANOR: Um….pracs.*
> *RESEARCHER: So you like doing practicals….why's that?*
> *ELEANOR: Because they're fun to learn and fun to see how they turn out.*
> *RESEARCHER: Okay, what do you find least enjoyable when we're in science?*
> *ELEANOR: Um…*
> *RESEARCHER: Do you ever find yourself getting bored?*
> *ELEANOR: Sometimes maybe. When we talking about things that, like, and words and stuff that I don't know. Like, really big words*
> *RESEARCHER: Really big words…do you find them hard to understand what they mean? Or?*
> *ELEANOR: Yeah.*

Eleanor made a link between vocabulary and her enjoyment of the class, but the problem was a little more than that. Referring to the International Competition for Australasian Schools test performed earlier in the year (of the research project), she insinuates that vocabulary often stops her from accessing many of the questions:

> *ELEANOR: Um.. no really only in some tests where sometimes things don't really make sense. Like on the ICAS I didn't get much of the questions.*

RESEARCHER: So do you sometimes find understanding what the question is asking is difficult?
ELEANOR: Yeah
RESEARCHER: Or is it that you're not sure about the concept?
ELEANOR: Well I understand what the question is saying but I just don't know what to do about it….

Eleanor is not alone in this regard. Peter is a student who sits roughly in the middle of the class in terms of academic progress and his writing skill. After completing the intervention process, he still makes it clear, that even though he knows the basic structure and requirements for the writing, he finds it hard to use the appropriate vocabulary:

RESEARCHER: What do you find least enjoyable about science?
PETER: Complicated words like the ones I can't really pronounce or spell.
RESEARCHER: What about what the words mean? Do you find that hard at all?
PETER: Yeah, I find that hard as well.
RESEARCHER: So, does that mean you don't find it enjoyable when you don't know that?
PETER: No, I do find it enjoyable, but sometimes I just get tripped up on words and stuff like that.
RESEARCHER: Maybe enjoyable is not the right word. Would frustrated be a better word?
PETER: It's only with some words…not all the time.

And then later in the interview...:

RESEARCHER: Do you find it hard to know what to write?
PETER: Sometimes, I get a little bit lost for words…what to write about.
RESEARCHER: When you say 'lost for words' do you mean that you're not sure what you have to write or that you're not sure what words to use.

PETER: Yeah, what words to use. If you said when you write a conclusion about the hypothesis I know what to do, but sometimes I just can't figure out which word to use and which one's better to use.

Eleanor and Peters' frustration with the jargon of science also contributes to their motivation and enjoyment of the subject. This is consistent with Hohenshell, Woller and Wallace (2013) who describe a student's view of vocabulary as a difficult landscape (p. 38). Bradley makes direct links between his enjoyment of science and the technical nature of science vocabulary:

RESEARCHER: Okay. Do you enjoy science?
BRADLEY: Yeah.
RESEARCHER: What do you find enjoyable?
BRADLEY: The experiments and, like, when we went outside and made the thing blow up.
RESEARCHER: So going outside to make things blow up yeah?
BRADLEY: (nods)
RESEARCHER: So you like seeing things that are 'spectacular'?
BRADLEY: Yeah
RESEARCHER: Okay. What do you find least enjoyable about science?
BRADLEY: Probably the reflections or something…. Like writing heaps of stuff.
RESEARCHER: So you don't find writing enjoyable?
BRADLEY: I like writing stories…
RESEARCHER: You like writing stories?
BRADLEY: Yeah… yeah.
RESEARCHER: What is it, do you think, about writing in science that maybe turns you off a little bit?
BRADLEY: Its just, like, at my old school we didn't do science, so I don't know anything about….like….m…c whatever….like all the names….
RESEARCHER: What do you mean by 'm' and 'c'?
BRADLEY: Like, Hydrogen's 'H' and the symbols.
RESEARCHER: Ahh, the symbols…
BRADLEY: Yeah…

Interestingly Bradley is not adverse to writing, he claims that the technical nature of science and lack of creativity when compared to writing 'stories' are what make the science writing less enjoyable. Proving that the opposite was also true, Arthur takes a different standpoint from Bradley and Peter:

> *RESEARCHER: It's almost like we're using a different language in science, do you find that difficult at all?*
> *ARTHUR: Not really*
> *RESEARCHER: No? Okay , what do you find most enjoyable about science?*
> *ARTHUR: Um.. all of the facts and experiments are fun also and you're just learning new information*
> *RESEARCHER: Is there anything about science that you don't enjoy very much*
> *ARTHUR: Ah no…*
> *RESEARCHER: Well that's nice to hear! Not even boring at times?*
> *ARTHUR: No.*

It is unsurprising that his positive attitude can be directly linked to his understanding of the science vocabulary and consequent success in classroom activities and assessments.

4.2.3 Theme 3 – Enjoyment and Success

In fact, enjoyment and success seem like intertwined concepts that permeated the students' responses throughout the interviews. Jenny was in the top five students in the class in terms of prior assessment of student achievement and made the following observations during her initial interview:

> *RESEARCHER: How do you view yourself as a science student, a student of science?*
> *JENNY: Um, I'm not sure actually….I think I'm okay at science.*
> *RESEARCHER: So you don't find it very difficult?*
> *JENNY: No.*

RESEARCHER: So you feel like you're good at science….why is that? How do you know?
JENNY: Well, its just… I enjoy it, so I feel like you do better when you enjoy something.
RESEARCHER: Okay. What do you find enjoyable about science?
JENNY: Well… I like how we get to do experiments and we can talk about different things… and we…I'm not sure…..hmmmm…
RESEARCHER: That's fine, you've said that you like experiments and that's good…..Is there anything about science that you don't enjoy so much or maybe that you find a little boring? That you're not so enthusiastic about.
JENNY: No, not really

Like Arthur, Jenny had experience with success in Science and consequently felt enthusiastic about the subject. She expresses it the other way around, saying that: "I feel like you do better when you enjoy something".

The successes that most of the students enjoyed over the course of the writing intervention project started to filter through into their own personal view of themselves. In David's final interview he explained how he felt more 'with-it' in terms of his writing:

RESEARCHER: Yeah? Do you think your writing is of a higher standard than what it once was?
DAVID: Yes, because in first term we only knew one acronym and now we know, like, a few. And I've got the hang of it.
RESEARCHER: How does the acronym help you?
DAVID: Going like step by step instructions. So, like, if we had CRABS I would just go C, R, A …. In that order.
RESEARCHER: And that tells you…..what? What does that tell you to do?
DAVID: Um…like, which step to go first.
RESEARCHER: So is that how you 'structure' a piece of writing.
DAVID: Yeah.
RESEARCHER: How does that make you feel about writing in science?
DAVID: Um…It makes me feel a bit more like…'in science', than before

RESEARCHER: Why do you think that is?
DAVID: Because I know more about it….yeah. [Smiling]

The smile on David's face was apparent to the researcher as a sign that he was feeling good about his improvement. Some were a little less committal to the idea. Lucy was a mid-range student in terms of her academic work, but her writing was of a high standard. After completing the intervention strategies with the rest of the class her responses in the final interview seemed to be tarnished with some lack-lustre performance on a recent class test:

RESEARCHER: You're finding the conclusions easier?
LUCY: Yeah, like understanding the topic….
RESEARCHER: Is there anything that you're finding challenging still?
LUCY: Um… kind of…..its hard to explain….some of the later questions in tests….
RESEARCHER: Some of the more difficult questions?
LUCY: Yeah

And then later in the interview…:

RESEARCHER: Do you think that when you do your writing now that it is of a high standard?
LUCY: Yeah, some of the time. Sometimes when it's a harder topic it's a bit harder.
RESEARCHER: So if I asked you to write a conclusion about an experiment that you'd be able to write a pretty good one?
LUCY: Yeah…..but it depends on the experiment.
RESEARCHER: And how do you feel about that?
LUCY: Well, I don't know really.
RESEARCHER: Well, do you feel better about yourself, to know that your writing is of a higher standard?
LUCY: Yeah.

There are direct links between success and motivation according to the literature. As Bandura (1994) points out, self-efficacy can be improved by experiences of mastery. In all of the interviews conducted with the students, it was evident that those who felt successful on tasks performed in class exhibited signs that it was an enjoyable experience and conversely those who didn't found it boring and unenjoyable. This also contributes to their perceived mindset (Dweck, 2000) as either 'fixed' or 'growth' and the implications of these mindsets as described in Section 2.5.

4.2.4 Theme 4 - Support for F.R.E.S.H Acronym

In most of the interviews conducted after the intervention strategies had been conducted, a common theme supporting the use of the scaffold as the best way to improve student writing was evident. The students found that the structure of F.R.E.S.H was easily remembered using the mnemonic device and that it made their conclusions more coherent and logical. Lucy identified that quantity did not always mean quality and that she found it easier to use science vocabulary in the right context:

> *RESEARCHER: We're going to move now onto how you write in science. So if you remember, we spent a lot of time looking at the FRESH acronym…*
> *LUCY: Yeah*
> *RESEARCHER: …and we looked at good examples and bad examples. We tried to peer assess each other, so we could see who was doing the right things, and we were also writing in our journals…*
> *LUCY: Yeah*
> *RESEARCHER: …and we showed examples of how we were getting better and we thought about what sorts of things were helping us to get better? So do you think those things have helped you all?*
> *LUCY: The acronyms have helped… of those…*
> *RESEARCHER: The acronym in particular?*
> *LUCY: Yeah.*
> *RESEARCHER: So the FRESH acronym, why is that one more helpful than the others?*
> *LUCY: Just remembering how to structure a conclusion. Not just writing on…*

RESEARCHER: What makes the structure of the conclusion….better, do you think?
LUCY: You don't include unneeded explanations, you don't….. [words indecipherable]
RESEARCHER: Would you say its fair to say that some people 'waffle on'
LUCY: Yeah [laughs]

And later in the interview…:

RESEARCHER: How do you know that your writing is at a higher standard?
LUCY: I'm using more complex vocabulary…and I'm using more information…. like whether the experiment worked and if the hypothesis was correct and stuff like that.

Similar to Lucy, Bradley also elucidated the fact that his writing had improved. He recognised that parts of his conclusion were either in the wrong order or simply missing at the beginning of the process:

RESEARCHER: So you're talking about the actual use of the FRESH acronym? So that actually has helped you?
BRADLEY: Yeah.
RESEARCHER: So do you think your writing is of a higher standard than it was?
BRADLEY: Yeah.
RESEARCHER: Okay, what's different about it do you think?
BRADLEY: Like, I get it in the right order and stuff like that.
RESEARCHER: Yep, what about actually 'what' you are writing. Do you think that's any different?
BRADLEY: Mmmm….what do you mean?
RESEARCHER: Well, before when you wrote a conclusion would you have always put those things in do you think? The things in FRESH?
BRADLEY: Umm…no……no.
RESEARCHER: So would you say that's an improvement?
BRADLEY: Yeah.

Karen was a girl whose prior assessments of achievement placed her in the middle to low ranking of the class. She recognised that the amount that she was writing was greater than it used to be and that the substance of the writing had also improved. She attributes this mainly to the F.R.E.S.H. scaffold:

> *RESEARCHER: Okay. Let's talk about how you've been going in terms of your writing. Um we've been using things like FRESH acronym, we've looked at um – you know – examples that are good or bad, and we've actually peer assessed each other about how people are writing and we've also been writing in our journals and thinking about what's a good strategy and what doesn't work so well. Do you think those things have helped?*
> *KAREN: Yeah, at the start I really couldn't think… but now it sort of takes me a long time to write it but its really good.*
> *RESEARCHER: Yeah?*
> *KAREN: Its better than it used to be.*
> *RESEARCHER: So you think that you've definitely improved?*
> *KAREN: Yep.*
> *RESEARCHER: That's good! Alright which one do you think has been the most help to you to actually improve your writing?*
> *KAREN: Um…..probably FRESH*
> *RESEARCHER: Ok, so using the acronym?*
> *KAREN: Yeah*
> *RESEARCHER: So what is it about that, that actually is useful for you?*
> *KAREN: Um….i don't know….it just sorta helps me know what it is I have to do, instead of just saying 'you have to write a conclusion or an introduction'*
> *RESEARCHER: Yeah*
> *KAREN: So it's easier for me to know what I have to go through*
> *RESEARCHER: Alright so, now do you think you're writing at a high standard in science?*
> *KAREN: Higher than I was, yeah. Much higher.*

Like the majority of the students interviewed, Karen describes writing as 'easier' after used the FRESH scaffold. Karen also realises that she is writing at a better standard even though there is still room for her to improve.

4.2.5 Theme 5 - Support For Peer Assessments

Two of the students interviewed, Arthur and Eleanor, reflected on the invention process and mentioned that peer-assessment had been beneficial to them in terms of learning about what to do and what not to do in the writing of their conclusions. The process of critiquing others was found to help them recognise the parts that they themselves did well and recognise the flaws in their work:

> *RESEARCHER: Sorry, maybe I should clarify. I meant in terms of the ways that we learning about writing, did you find any of the methods more useful you to learn how to write a good conclusion?*
> *ARTHUR: Well I found the peer and self-assessment [sic] of the practise tasks to be the most useful. And maybe looking at good examples and bad examples….*
> *RESEARCHER: What was it about looking at somebody else's work that helped you out do you think?*
> *ARTHUR: Because we sort of know where we are around…*
> *RESEARCHER: Sorry, I'm not sure I understand….*
> *ARTHUR: Um…..I don't know.*
> *RESEARCHER: Let's say I gave you somebody else's work and you peer assessed it. How does that help you to improve your own work?*
> *ARTHUR: Ohh, because we have to correct their work to improve it and so we can do that with our own work as well and, sort of, improve it…. And improve their work.*
> *RESEARCHER: Do you think the phrase learning from other peoples mistakes is applicable here?*
> *ARTHUR: Yeah.*

Eleanor expressed the same sort of opinions as Arthur, valuing the peer assessment as a different way to receive feedback:

RESEARCHER: Is there one that you thought that was most useful. Out the peer assessment, the FRESH, and journaling….were any of them more useful than the others?
ELEANOR: Well, I did like the journaling, because it was, like, different from what we normally do. And I guess…didn't we do peer assessment after we wrote in the journal?
RESEARCHER: Yeah.
ELEANOR: Well, um. I think it was just easier because you got to write down the acronyms and then write from them, write what you thought.
RESEARCHER: Do you think that looking at other peoples work was helpful?
ELEANOR: Yeah because thinking of...like, knowing what they've written compared to yours… like, what the differences and how they're the same and their opinions of your work.

Eleanor's last comment in the quote shows that she was able to contrast her peers' work with her own and valued their points of view. Her positive response to journaling as something "different from what we normally do" seems to indicate that she also valued a different approach to learning to write conclusions in science. The ability to comment on others and think about their own work was quite novel and fresh for her.

4.2.6 Theme 6 – Metacognition Understated

In the interviews metacognition seemed to be downplayed in terms of its importance by the students, possibly because the interviewer mentioned it as 'journaling' without really explaining it. Jenny was one of the only students who seemed to make mention of a process of reflection in terms of 'doing science':

RESEARCHER: And we've put that up on the wall, and we also looked at examples where FRESH was being used and those sorts of things. We also looked at each other's work which we peer assessed a lot. And we also wrote a lot of things in the journal, to try and help us think about what we were doing and those sorts of things. Do you think they've helped you at all?
JENNY: Yeah.

RESEARCHER: Why's that?

JENNY: Well, it just gives me an easier…like….I remember when I used to be stuck on what to think and what to say but when we study it a bit more you, kind of, know more about it, so it's easier to do science theory and stuff like that.

RESEARCHER: Okay. Do you think your writing is of a higher standard than what it was?

JENNY: Um….yes…probably.

RESEARCHER: Why do you think that is?

JENNY: Well, just writing in general I've….its hard to explain but I've just practised writing more….so, like, when I'm writing stories I just write from my head, so its kind of like writing a conclusion because you don't really have…um….exact things….like….you need to talk about this, this and this….its kind of free for you to think what you want and write what you want.

Jenny recognises that she had trouble with particular elements of writing conclusions such as including scientific theory to support claims made about an experiment, for example. She knows that conclusions have specific 'parts' that make them of a high standard (she is probably referring to the 'parts' of the F.R.E.S.H. scaffold). She finds it hard to pinpoint a particular improvement strategy and stated "I've just practised writing more".

In terms of metacognition, the majority of students argue for the use of the F.R.E.S.H. acronym as a structuring device for their writing, whereas others tended to favour getting the opinions and feedback of their peers for improving their writing. Whether or not students' valued the process of metacognition as an improvement process in itself does not seem to have any direct evidence in terms of their responses in the interviews. As a high-order level of thinking, the process of metacognition may not have occupied their thoughts as an 'improvement strategy'.

4.3 Student Journals

The students' journals contained comments made by the students and their peers as well as the physical writing sample of conclusions that they wrote over the course of the intervention process. As in the previous section, the students' names have been changed to protect their identity and any pertinent information about their background is included as their comments appear in the discussion to maintain an authentic voice.

Section 4.3 has been organised into three parts that discuss use of the F.R.E.S.H. scaffold and modelling, peer/self-assessment and lastly, metacognition. Themes that emerge are discussed within each of these parts as required in a narrative style with student voices in the form of journal quotations forming the 'dialogue' of the discussion.

4.3.1 F.R.E.S.H. Scaffold and Modelling

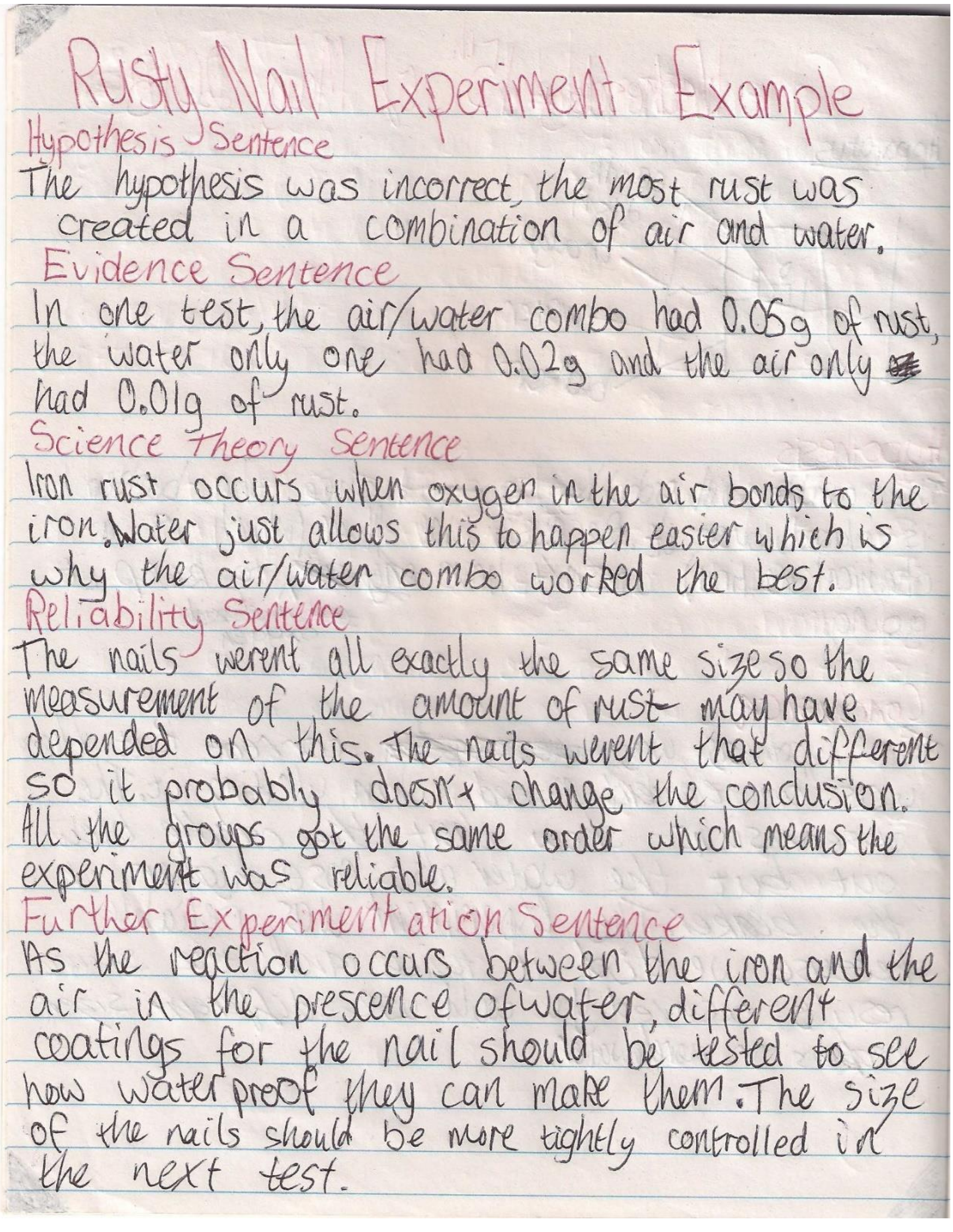
Figures 4.1 and 4.2 show two examples of the F.R.E.S.H. acronym being used as an organiser for a conclusion as headings or written in the margin of the student's page. Figure 4.1 was written entirely by Abigail as a response to a class experiment where nails were allowed to rust in different environments. The example in Figure 4.2 is from a co-constructed conclusion written (although incomplete) by Rowena as part of an entire class discussion of a 'strong' conclusion. Abigail and Rowena were two girls of similarly quiet demeanours who were in the upper middle and lower middle (respectively) of the class in terms of academic ranking.

Later, in the summative assessment for the skill of concluding, students were asked to access some secondary data and draw a conclusion from it. Figures 4.3 and 4.4 show the writing of students using the F.R.E.S.H scaffold to structure their writing.

Part of modelling the process involved the use of strong and weak samples based on the F.R.E.S.H. scaffold. Students commented on the provision of these samples and their perceived utility. Allison was a strong student in the Science class and she had excellent written skills. She emphasises the structural components of the F.R.E.S.H. acronym and was positive about the use of modelled writing provided to her:

From Allison’s Journal: “I think I have improved in writing conclusion. This is because I know the real structure of fresh. Yes it is worthwhile to have a strong and weak sample to learn to write because you can examine the difference and see the mistakes. You can also see what good writing looks like.”

Figure 4.1 ***F.R.E.S.H. Scaffold Example 1 (From Abigail’s Journal)***



Rusty Nail Experiment Example

Hypothesis Sentence

The hypothesis was incorrect, the most rust was created in a combination of air and water.

Evidence Sentence

In one test, the air/water combo had 0.05g of rust, the water only one had 0.02g and the air only had 0.01g of rust.

Science Theory Sentence

Iron rust occurs when oxygen in the air bonds to the iron. Water just allows this to happen easier which is why the air/water combo worked the best.

Reliability Sentence

The nails werent all exactly the same size so the measurement of the amount of rust may have depended on this. The nails werent that different so it probably doesn't change the conclusion. All the groups got the same order which means the experiment was reliable.

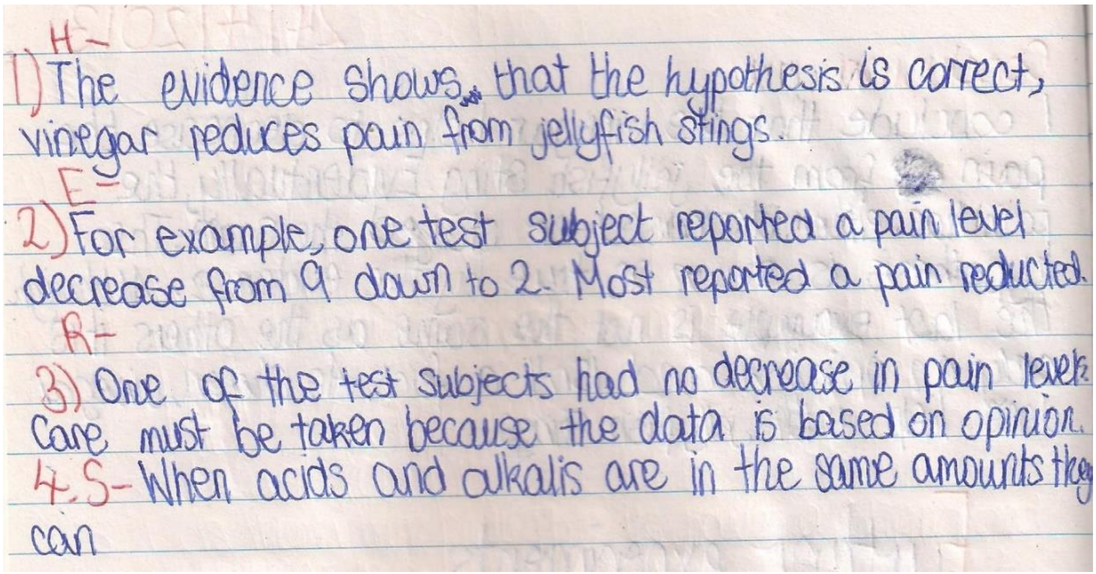
Further Experimentation Sentence

As the reaction occurs between the iron and the air in the prescence of water, different coatings for the nail should be tested to see how waterproof they can make them. The size of the nails should be more tightly controlled in the next test.

Steven, at the other end of the scale ability-wise, was in agreement:

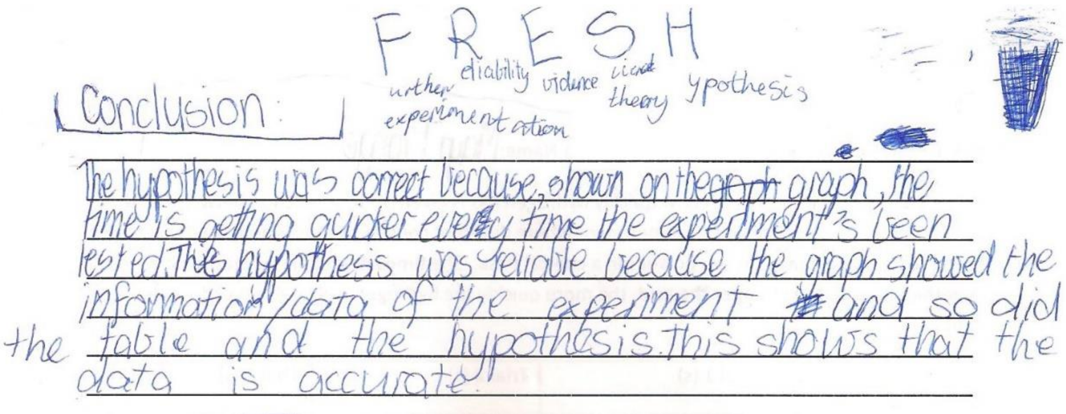
From Steven's Journal: "I think the strong/weak [sample] is a very good idea because it not only teaches us the good things to writ[e] but it teach[es us] what we shouldn't write."

Figure 4.2 ***F.R.E.S.H Scaffold Example 2 (From Rowena's Journal)***



H-
1) The evidence shows that the hypothesis is correct, vinegar reduces pain from jellyfish stings.
E-
2) For example, one test subject reported a pain level decrease from 9 down to 2. Most reported a pain reducted.
R-
3) One of the test subjects had no decrease in pain levels. Care must be taken because the data is based on opinion.
4. S- When acids and alkalis are in the same amounts they can

Figure 4.3 ***F.R.E.S.H Scaffold Used in Summative Assessment (Michelle)***



F R E S H
urther eliability vidence cientific theory ypothesis
experimentation

Conclusion:

The hypothesis was correct because, shown on the graph, the time is getting quicker every time the experiment's been tested. The hypothesis was reliable because the graph showed the information/data of the experiment and so did the table and the hypothesis. This shows that the data is accurate.

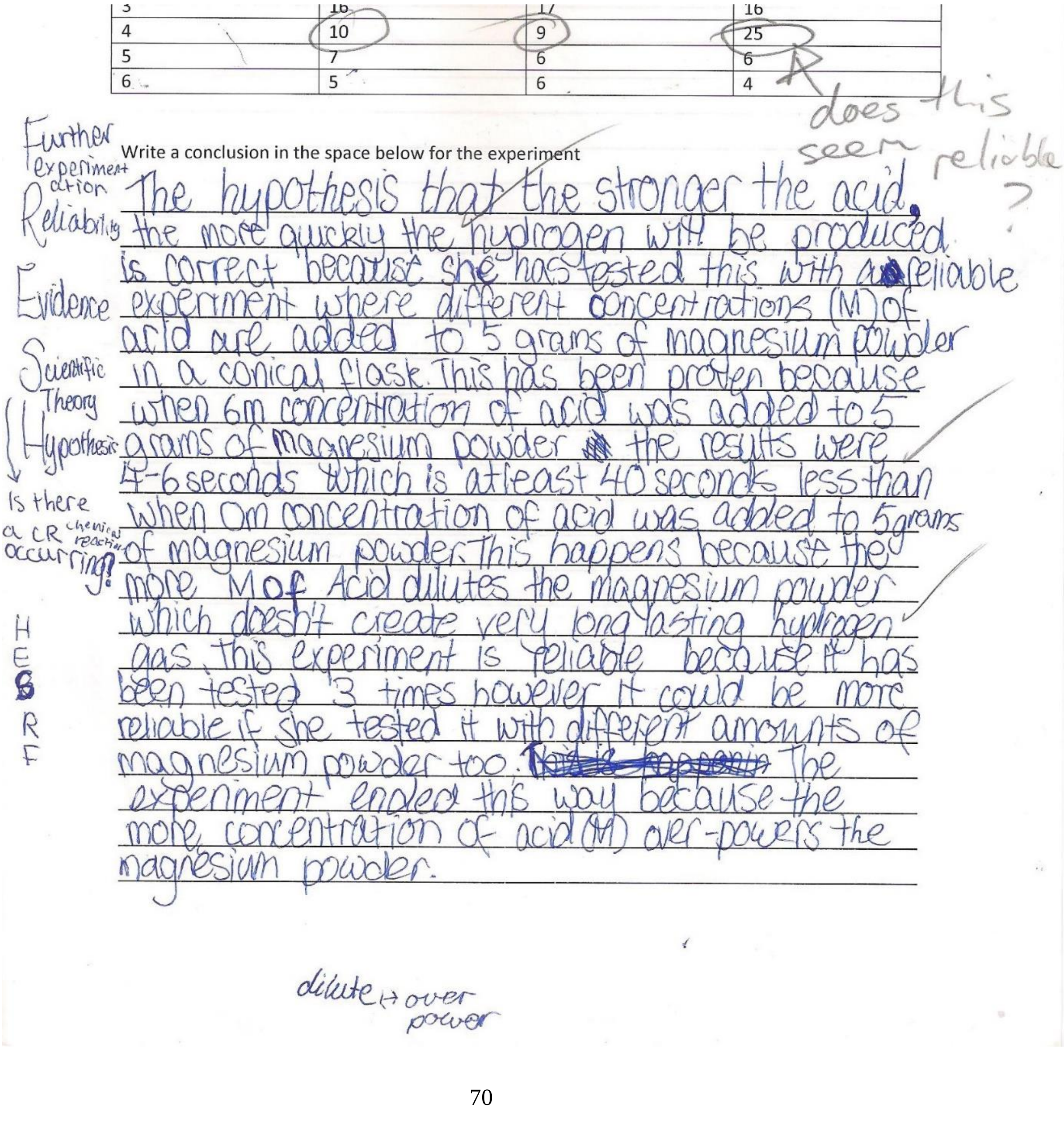
It is worth noting though that not every student was enamoured with the weak samples:

From Arthur's Journal: "the only unuseful[sic] strategy is seeing strong examples and weak examples because seeing weak examples doesn't help me in any way."

And also…:

From Abigail's Journal: It is not worthwhile to have a st[r]ong and weak sample because it is a waste of time. The weak samples are to[o] weak. (they sound like the[y] were written by 2 year olds!)

Figure 4.4 ***F.R.E.S.H Scaffold Used in Summative Assessment (Rowena)***



3	16	17	16
4	10	9	25
5	7	6	6
6	5	6	4

Write a conclusion in the space below for the experiment

The hypothesis that the stronger the acid, the more quickly the hydrogen will be produced is correct because she has tested this with a reliable experiment where different concentrations (M) of acid are added to 5 grams of magnesium powder in a conical flask. This has been proven because when 6m concentration of acid was added to 5 grams of magnesium powder the results were 4-6 seconds which is atleast 40 seconds less than when 0m concentration of acid was added to 5grams of magnesium powder. This happens because the more M of Acid dilutes the magnesium powder which doesn't create very long lasting hydrogen gas. this experiment is reliable because it has been tested 3 times however it could be more reliable if she tested it with different amounts of magnesium powder too. ~~This is happenin~~ The experiment ended this way because the more concentration of acid (M) over-powers the magnesium powder.

Nor did every student agree that the F.R.E.S.H. acronym was easy to use. Two of the students stated that as it was a 'reversed' acronym that confused them in terms of the order in which to write the appropriate parts. Nathan was strong academically but often had difficulty with writing which may have stemmed from English being a second language at home, although he had been immersed in the English language at school from a young age:

> From Nathan's Journal: "The weak and strong examples were useless and 'Fresh' is hard to remember. I often have to think FRESH as HESRF so I don't right [sic] the conclusion in the incorrect order."

And again from Steven…:

> From Steven's Journal: "The FRESH acronym I don't find very helpful because even though it says the parts of a conclusion its not really in order."

These students both seemed to prefer the peer-assessment feedback as a 'better' way to improve their writing.

4.3.2 Peer/Self-assessment

Initially the peer/self-assessment took the form of comments written about their own or others' work having been given ample time to read and absorb the writing. Initially, the feedback that students gave was brief and non-constructive. After describing the F.R.E.S.H. acronym to the class and asking them to think about a conclusion that they had written previously, Jenny wrote two words in her journal next to the writing - "very bad". In those early days of the intervention process, Jenny seemed more able to give advice to others:

> From Allison's Journal: "You could have put a few more different formulas like NaCl, HCl and NaOH. You might have put who you paired with and what you were. Good Job! From [Jenny]"

This was typical of the students' advice to each other, with some small bits of constructive criticism with (normally) some form of congratulation. Some other samples that show this are given in the quotes below:

"I think you could have written more about what we actually did. But apart from that it's brilliant! Good Job!"

"Well done [Nathan], you could have said more about how it neutralises each other but it was very accurate. Good work, [Tina]."

"Awesome job xoxoxo. Try next time to elaborate more on the main features: the symbols and substances."

The students were able, with very little prior training, to identify one feature of the writing that they felt was important for improving upon, but found it difficult to frame an improvement strategy for the peer to employ. They also felt that it was important to make their peer 'feel good' with some form of congratulatory comment.

It was at this point that the researcher realised that a more structured approach to the feedback was required (the reasoning for this is discussed in the next section). A rubric was co-created with the students to give them a framework for their feedback to each other (Refer to Table 3.5) and also to give a personal comment on a 'star' (one thing that they had done well) and a 'step' (the next thing that they should concentrate on improving).

Figures 4.5 and 4.6 are examples of collaborative peer feedback using the co-constructed rubric. This feedback was completed in groups of 2 to 3 students using the activity in appendix B.

The student response to the peer assessment was generally positive. After a few instances of its use, the students were asked to respond to the techniques that they had been using to improve their writing. Steven equated the feedback that the researcher had given him to feedback from his peers:

> From Steven’s journal: “Yes I have found these techniques very useful because peer assessment not only tells me what the teachers think but also what the students think.”

It is hard to tell whether he thought the feedback from his peers was as useful as that from his teacher but, as Tina points out, the peer feedback may have had some drawbacks. Tina was one of the top students in the class and had excellent writing skills. Her frustration with the peer feedback shows in the following excerpt from her journal:

> From Tina’s journal: “I have found the peer assessment good because I get feedback and comments I can improve on. In some ways it wasn’t good because if I were in a group with certain people their friends would just give us high marks and the person in our group only wanted to score them high. I found peer assessing interesting depending on who I was with.”

To reinforce Tina’s point, Michelle, a girl whose classroom behaviour seemed to revolve around social interaction with her peers, was very enthusiastic about peer assessment:

> From Michelle’s Journal: “I did like the peer assessment because I thought that it’s a great way to get assessed. You know what it feels like to have experience being judged on something by other people.”

Figure 4.5 ***Peer feedback Performed on 'Conclusion 1' of the Acid and Limestone Strong/Weak Sample Activity (Appendix B)***

Not at required level	At level	Above required level
Hypothesis evaluation not included	Hypothesis is evaluated as right/wrong	A general statement is used to describe the outcome of the experiment
Evidence is not used to support the findings	Observations or numerical data are used to support the findings	Multiple sets of data or observations are used to support the findings and demonstrate any relationships
Scientific theory is not mentioned	A brief explanation of reasons for the data is given	An in depth discussion of scientific theory is used to describe why particular data was obtained
Reliability is not mentioned	The reliability of the data/observations is discussed	The reliability of the data/observations is described in detail and as well as the effect of error on the conclusion's validity
Further experimentation is not mentioned	Suggestions that may further the understanding of the aim are included	Modifications to the method to improve validity are suggested as well as other variables to change

Comments

Star

Good science theory was strong.
Included all symbels. Explanation.

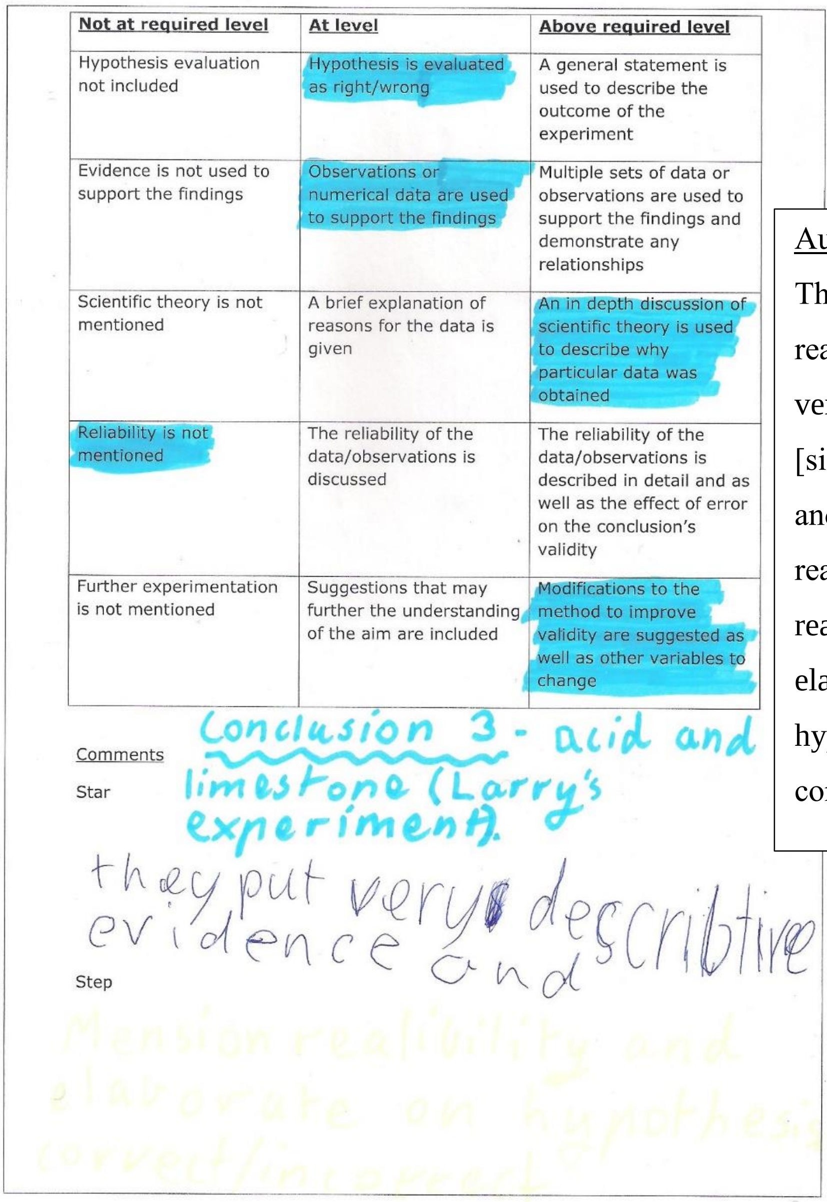
Step

Try to add more detail for evidence and reliability

Author's notes

The yellow highlighter used by the student has been enhanced by a digital circle placed around the area where the students indicated on the rubric

Figure 4.6 ***Peer Feedback Performed on 'Conclusion 3' of the Acid and Limestone Strong/Weak Sample Activity (Appendix B)***



Not at required level	At level	Above required level
Hypothesis evaluation not included	Hypothesis is evaluated as right/wrong	A general statement is used to describe the outcome of the experiment
Evidence is not used to support the findings	Observations or numerical data are used to support the findings	Multiple sets of data or observations are used to support the findings and demonstrate any relationships
Scientific theory is not mentioned	A brief explanation of reasons for the data is given	An in depth discussion of scientific theory is used to describe why particular data was obtained
Reliability is not mentioned	The reliability of the data/observations is discussed	The reliability of the data/observations is described in detail and as well as the effect of error on the conclusion's validity
Further experimentation is not mentioned	Suggestions that may further the understanding of the aim are included	Modifications to the method to improve validity are suggested as well as other variables to change

Comments

Conclusion 3 - acid and limestone (Larry's experiment).

Star

they put very describtive evidence and

Step

Mension realibility and elaborate on hypothesis correct/incorrect

Author's notes

The 'star' part reads: "they put very describtive [sic] evidence and" and the 'step' part reads: "Mension realibility [sic] and elaborate on hypothesis correct/incorrect"

Some students did not like the experience of peer-assessment. One girl in particular had difficulty with the process:

> From Beth's Journal: "I do not like peer assessment either because it is really hard for me to mark someone else's work"

It is hard to tell from her comment whether she means that it is technically hard to mark her peers work as she feels inadequately skilled or whether she feels that it is socially hard to criticise her friends. From a background perspective, Beth seemed to be a very introverted girl, whose prior assessment placed her in the lower middle of the class academically speaking. Unfortunately this aspect was not investigated any further during the research.

4.3.3 Metacognition

In terms of metacognition, the students identified either the F.R.E.S.H. scaffold or asking for feedback in terms of Peer Assessment as the best ways to improve their writing. Many students recognised that the structuring element of the F.R.E.S.H. acronym gave them 'topic sentences' that were important to the overall utility of the conclusion:

> From Michelle's Journal: "I find FRESH a great way to write up a hypothesis because when you use it you're writing everything you need to include except you're elaborating."

> From Allison's Journal: "Peer assessment and the examples were not as useful as FRESH because I wouldn't know what elements I need to include in my writing."

Some students recognised that, when they didn't use a strategy to help them write, it resulted in a less coherent piece:

> From Beth's Journal: "I only think I improved a bit. Because I didn't really use the FRESH acronym so I didn't improve much!!"

Some students persisted with a modified 'more is better' theme. When students were only writing one or two sentences about the outcome of the experiment, they missed vital parts of the conclusion altogether. It is worth highlighting the fact that the F.R.E.S.H. scaffold indicates that a minimum of five separate points (one for each part of the acronym) are required for a conclusion to be considered 'successful'. In order to achieve higher levels on the Conclusion Writing Continuum (Table 3.6), students must provide more detail and consideration for each of those five areas. The 'more is better' assumption is an astute and true observation made by the students in this regard. Even so, the substance of the writing is far more critical than the quantity. A student could potentially write pages about what they saw happen in an experiment without suggesting a single reason why it happened. Alternatively another student could use succinct and efficient phrases to fit the entire conclusion into the size of a single paragraph:

> From Jenny's Journal: "I think I have improved in making conclusions for at the start I had trouble thinking of what to write and the second one was ended though I had more to write."

The second part of Jenny's sentence is probably referring to the fact that she was running out of time and could have written more in her conclusion. She had recognised that parts were still required to complete a 'good' piece. In the interests of internal validity, it is recognised that she could be referring to quantity alone (regardless of quality/content) as the defining indicator her improvement. Given that the statement was made directly after the class had been introduced to the F.R.E.S.H. acronym as the guiding principle for structuring their writing, this is deemed unlikely.

Another theme that emerged was the equating of the strategy to enjoyment. Students are more likely to engage with the strategy with which they feel most comfortable or enjoy the most:

> From Lucy's Journal: "I found only FRESH useful because it did help me write a more expanded conclusion which made mine better in explaining the parts of a conclusion easier. I don't like the others because I find them really boring."

The enjoyment may come from the fact that she had more success with that strategy and felt better about herself as a result. If the other strategy was perceived as giving her less gain, she may decide that they were of little worth and therefore "boring".

Some students recognised the utility of both writing strategies and were amenable to both:

> From Arthur's Journal: "Most of these strategies of learning how to write a decent conclusion are successful in my writing. The FRESH acronym helps me to remember all of the crucial parts of a conclusion I must include. Peer assessment also helps me to become stronger at writing conclusions because I can see different conclusions."

Other students definitely had a preference for just one of the strategies:

> From Nathan's Journal: "I found peer assessment a better technique as it was like easier to remember. Normally I remember what other people say rather than FRESH and examples."

Overall, the student analysis of the strategies that they had practised was limited to simple statements that summarised the overall worth of the strategy to their level of improvement.

One student, Ralph, had quite a negative view of himself and, whilst this is not exactly metacognition, his comment shows his reflection on his own work and consequent attitude towards Science:

> From Ralph's Journal: "I think I have drop [sic], I just do not understand Science. It is Just does not like me. [sic]"

This was in response to his self-assessment of his first conclusion after introduction to the F.R.E.S.H scaffold and listening to what others had written. He recognised that his work was at the lowest possible level and his perception of the amount of improvement required was probably very daunting for him. It should be noted that this was the only comment of its kind in his journal, and that it was made very early in the intervention process. His subsequent writing improved and his personal comments were limited to a description of the improvement strategies that he used in class.

4.4 Teacher Journal

The teacher/researcher kept an electronic journal over the course of the implementation of the intervention strategies. Most of the entries represent reflections and observations made after most classes where the strategies were used and practised, although some entries are thoughts about the overall progress of the study. Much of the journal is presented in the first person and it seems logical to present the excerpts chronologically as a narrative of the intervention process. Discussion of themes are presented as they occurred and many represent rationale for the use of various parts of the intervention as part of the 'feedback-loop'. In terms of internal validity (Polkinghorne, 2007, pp. 475-476, Quoting Campbell & Stanley, 1963), as this represents the researchers own views, it is apparent that there would be little to no misinterpretation of the data in the journal in terms of intent or causation.

The first class during the intervention phase was designed to get a benchmark for the initial ability levels of the students given a set of second hand (and fictitious) data. Analysis of this data is addressed in the next section on quantitative results. Also, the students got their first taste of peer feedback in that lesson. The reflections in the teacher journal were arranged sequentially and were written directly after each class that had produced data for the research project:

> *From Teacher Journal: "Initial class"*
>
> *"I gave the class a set of artificial data and asked them to tell me what happened and what the experiment showed, 'Write a conclusion.'*

> A chemist wants to know if adding vinegar to a jellyfish sting is beneficial. She hypothesises that jellyfish stings are alkaline and that vinegar should neutralise it. (Actually we're not sure why it works!)

Pain reported on a scale of 1-10 of a jellyfish sting	
Before applying vinegar	After applying vinegar
5	3
7	2
9	2
10	6
8	8

The students read their conclusions to each other - surprisingly some brave volunteers who prefaced their reading with "I don't think I did a good job" were quite keen to read theirs out to get some feedback.

Students identified what was good about the conclusions in a class discussion. Some comments were "it was quite clear", "I don't think they did a great explaining WHY it happened", "I liked the suggestion about testing different jellyfish" "I wonder why one of them didn't think it did anything?""

The researcher was genuinely surprised at how willing some of the students were to get feedback from their peers, even if they had already self-assessed themselves negatively. He had a preconceived notion that the students would be reluctant to offer their work up for critique due to a need for acceptance among their peers. Allowing other adolescents to say negative comments about a piece of work that they had created seemed like the last thing a Year 6 student would want. The researcher believes that the students felt safe enough with their peer group to be more adventurous with their learning and not to fear 'put-downs' or shame.

The class discussion of the initial pieces of writing shows that students were able to comment on aspects such as clarity and content. From the nature of the quotes recalled by the researcher after the lesson, it seemed as if the students were not able to give much description on how to improve the pieces. Their comments were limited to what was 'liked' or 'disliked' and other non-constructive evaluations.

The second part of the introductory class was devoted to unveiling the F.R.E.S.H acronym with which the students would become very familiar over the next few weeks. The researcher and the students used the scaffold to 're-write' the conclusion collaboratively during a discussion in which the students' ideas and words were used to write the model example:

From Teacher Journal: "Initial class"

"Introduced students to FRESH acronym today – some had seen it before
Further experimentation
Reliability of data
Evidence that shows what happened
Scientific theory that could explain the results
Hypothesis correct or not.

We then co-constructed a conclusion using the criteria

1) *H- The evidence shows that the hypothesis is correct, vinegar reduces the pain from jellyfish stings.*
2) *E- For example, one test subject reported a pain level decrease from 9 down to 2. Most reported a pain reduction.*
3) *R- One of the test subjects had no decrease in pain levels. Care must be taken because the data is based on opinion.*
4) *S – When acids and alkalis are in the same amounts they can neutralise each other which might be why the pain was reduced.*
5) *F – To test this further, the stings could be tested in a laboratory rather than using people's opinion.*

Students seemed to accept that our co-construction was superior to what they had already written."

The co-constructed sample shows the use of the letters in F.R.E.S.H written as a structural organiser for the conclusion which some of the students would later employ to help write their own pieces. Following the introduction to the F.R.E.S.H. scaffold, the student were then asked to try writing another conclusion with another set of second-hand (and fictional) data:

From Teacher Journal: "Initial class"

"I gave them a second set of data to try a conclusion

A chemist wants to know if solution X is an acid or an alkali. He hypothesises that it is an alkali because it was found with other cleaning products (which are normally alkalis). He puts a drop of universal indicator into the solution and records the colour firstly as 10mL of hydrochloric acid is added, and then secondly as 10mL of NaOH (sodium hydroxide base)is added.

Amount added (mL)	Colour with HCl acid	Colour with NaOH base
0	Yellow	Yellow
10	Orange	Green
20	Orange/Red	Green/Blue
30	Red	Blue
40	Red	Blue/Purple

The students were starting to grow restless. Some students seemed to have had enough of writing. Maybe I need to attack this a little more slowly –rather than spend an entire period on it. Some of the students with shorter attention spans might start disengaging.

It would have been good to do another peer assessment but we ran out of time so I asked them to reflect on what they had done – had they improved? They wrote this in their learning journals. From their restless behaviour I gathered that more than a few were sick of writing by the end! Overall I believe we had made some headway in this class – the students seemed to genuinely recognise that a conclusion was more than just a single basic statement about what happened. The biggest challenge seems to be getting the association between *the theory behind the experiment and evidence that supports it. My fear is that if they don't get the theory how will they be able to communicate it in their writing?"*

By the end of the lesson, the researcher had noted a general improvement in the standard of the science writing performed by the students. It was becoming evident, though, that time was going to be an issue for the research in two main areas. Firstly, the students were becoming "restless" and "sick of writing" and found it hard to remain focussed on the 'improvement of writing' for an entire hour-long class. The implementation of the writing improvement strategies would need to be done in a more 'piecemeal' approach with no more than half an hour dedicated to it during any one class. Secondly, the process of writing took a lot more time than was originally thought. With more to write about, naturally the writing would take longer. It was noticed that some of the students were unable to finish in the allotted time. Sometimes this was due to poor handwriting skills or inefficient language skills.

Another important point made in the journal entry above was the observation about scientific 'theory'. The students had just finished learning about Acid/Alkali neutralisation reactions and the uses of universal indicator. The 'S' in F.R.E.S.H. is to get students to consider why the experiment yielded the results but it did and to make an inference based on scientific theory. Student understanding of the scientific theory may block the student's ability to write about it. This fed back into the teaching of the strategy and accompanying discussions of experiments performed in subsequent classes. Many experiments were followed up by a question akin to "do you see how this data supports what we have been learning about?" or "which theory would best explain these results?" in order to encourage the process required to write the 'S' sentence in a conclusion.

With approximately one to one-and-a-half weeks between classes devoted to the research, some reminders about the preceding class was often necessary to keep the students 'up to speed' and to allow absent students a chance to catch up:

From Teacher Journal: "2nd class"

> *"Had students write a conclusion using FRESH – one sentence at a time. We started with Hypothesis sentence, then evidence supporting, reliability and science theory. We had to readdress what each part meant. I might make a poster out of the main steps. The further experimentation step was interesting because next lesson the plan is to get them to carry their further experimentation step. I think it is important to make the steps real – rather than make the F part some surreal thing that they will never do. I asked the students to read each other's work and write a comment, on circulating around the room I realised that most of the comments were 'great job' and non-critical. A few of them were focussed on 'more detail' type comments without being very specific. They will need to be very structured I think."*

The researcher makes note of the need for authentic experiences when writing. Butler and Nesbit (2008) made use of publishing students' work in a public forum as an authentic motivator. In the quote above, the use of students' ideas to inform the next experiment is noted as a "real" experience and gives purpose to the writing. The benefit of using first-hand data (generated by the students) over second-hand data (generated fictitiously or taken from a different source) is that the students feel ownership of the data. It is also assumed that they would be more likely to be successful when evaluating it in terms of errors made and their consequent effects on the data.

At times motivation was an issue. Sometimes external factors such as classroom environment (for example, high temperatures) and time of day could influence the overall 'mood' of the class. A lack-lustre response to a type of classroom activity might also be attributed to boredom with a repeated activity. Some modification to an existing activity to increase engagement was sometimes required as the following excerpt indicates:

From Teacher Journal: "3rd class"

> *"I had the students write a conclusion together with each other. They will hopefully be learning from their peers about what is important in a conclusion. Pointed out the fresh acronym but let them go for themselves. Some students were reluctant to write, citing 'I've got poor handwriting' as an excuse. My suspicion was more that they didn't want to. I forced the writer to change over so that everyone wrote something. The focus of a few distracted groups seemed to improve once they all realised they would be contributing."*

In this case, a small-group co-construction of a piece of scientific writing was utilised to encourage peer learning. It was noticed that some students were inclined to let 'the best writer' do the majority of the work and to disengage from the thinking involved with writing the conclusion. The researcher's response was to force a change in roles within the small group so that all of the students were engaged with the process of co-construction.

In the early stages of the research, the students' lack of experience with peer/self-assessment was showing. The observations of non-constructive feedback and the lack of specificity with regard to the few instances of criticism required a change to the way in which the peer/self-assessment was to be conducted:

From Teacher Journal: "3rd class – further reflection"

> *"After reading some of the journals tonight, I realise that these kids have no idea how to assess either their own work or each other's. I'm not sure they'll be able to be critical (constructively) of themselves if they can't be critical of each other first. I think they need some real structuring of their critiquing with a rubric or something or the 'star and step' method."*

The use of a rubric to structure the students' critiques of each other and also of themselves was deemed the best response to the inexperience with this form of feedback. In the interests of 'student ownership' the rubric was to be created with student language and responses to the structure of a conclusion based on the F.R.E.S.H. scaffold:

> *From Teacher Journal: "4th class"*
>
> *"We spent some time (20mins) today discussing how we might assess other peoples work or our own. I got students to say what they thought should go into the conclusion and kind of manipulated/paraphrased what they said live straight into a table. A few times I had to prompt them to think about the things that we had discussed with the FRESH poster like what the effect of errors are on the validity of the conclusion. I found it hard to get everyone saying something but generally everyone seemed engaged. I promised that I'd polish what we discussed and we'd only use it with our own class."*

On reflection, the researcher regrets the fact that the rubric only addressed the components of the F.R.E.S.H scaffold and paid no heed to overall cohesion and clarity of the writing as Butler and Nesbit (2008, p. 139) had done in their research.

To practise using the rubric, a set of conclusions was written for the 'Acid and Limestone' activity as shown in Appendix B. Some samples of student use of the rubrics are given in Figures 4.3 and 4.4. The improvement in the students' ability to give feedback seems apparent in the following excerpt from the Researcher's journal:

> *From Teacher Journal: "5th class"*
>
> *"I had the students look at some conclusions that I 'made up' about a set of data. I made them a weak, medium and strong sample. They used the rubric from last time to assess each of the conclusions. Strangely a number of them didn't like being saddled with the 'dumb' conclusion as though they thought that I thought it was all they could handle. I got each group to tell the class what they thought of each conclusion and how it rated on the rubric. I also asked them to give a star (what they did well) and a step (what they need to focus on) for the fictitious writer. From their discussion I can see that it is starting to sink in what is required for each part of the acronym. Hopefully*

they'll be writing some better ones soon. I asked them to tell me what they thought of the samples in their journals."

The researcher noted how poorly some students received the weakest of the three samples. It was as though it was a reflection on their own level of ability and some reassurance had to be given to make the purpose of the activity clear. Some of the more capable students questioned how valid the activity was in terms of the weak samples and they didn't believe that they taught them anything.

The introduction of the rubric improved the level of the student feedback to each other. It seemed that the level of feedback was proportional to the ability level of the students giving the feedback:

From Teacher Journal: "6th class"

"After today's burning candle experiment I asked the students to write a conclusion referring to the hypothesis they wrote at the start. I had them peer assess themselves using the rubric. They seemed very excited to get their feedback from their mates, but I got the feeling that some of the girls especially weren't taking it overly seriously. The comments are getting better – but still non-specific. My guess is that if they don't know themselves, they find it hard to do it for the others. I'd say the best feedback is coming from the strong students. I feel a little sorry for the kids who are being assessed by the weaker ones…."

The researcher makes a general observation that the best written feedback is given by students who are capable of writing that is better than the piece they are assessing. It is worth noting that the majority of students were able to use the rubric to show where the conclusion sits in terms of the parts of the F.R.E.S.H. scaffold, but their descriptions of the improvements needed were still quite vague.

The excerpt above also describes the student's reaction to the peer feedback process as "excited". Some of the girls in the class saw it as a social opportunity with each other and the validity of their assessment is called into question. Some frustration was experienced by the researcher when a couple of girls gave each other high scores which seemed highly unlikely given prior experience with their work. The students' lack of maturity in this case may have influenced their ability to separate the supportive niceties of friendship from their ability to indicate a true measure of each other's work.

Reacting to a desire to vary the teaching/modelling of the F.R.E.S.H. scaffold and the students' ability to assess writing of different ability levels, the Rusty Nail activity was given to the students (See Appendix B):

> *From Teacher Journal: "7th class"*
>
> *"I asked half of the students to write a strong conclusion based on their selection of the best sentences from the 'rusty nail' experiment activity and the other half to write out a weak one based on the worst sentences. A few volunteered to read out their strong and weak samples. One said 'I don't get why we are looking at weak samples – we only want to get better' to which I replied 'but this way you know what not to do! His body language implied that he didn't accept my reasoning – like I was wasting his time. My thoughts are that they probably don't see the value of it but I kinda [sic] saw his point."*

Again the worth of weak samples was called into question by the students. The researcher deemed the modelling of a weak sample to be important so that students could contrast it with a stronger one and recognise the aspects of both. This experience would help them to identify strengths and weaknesses in their peers' and their own work. It was hard to get the students to understand this line of thinking and the provision of weak samples was a source of dissatisfaction for many of them throughout the intervention process.

Towards the end of the in-class section of the research project, it was observed that some students were making a lot of progress, while others were still experiencing difficulty. Some negative attitudes towards writing were apparent in classroom observations but neither appeared (very often) in their journals, nor were displayed by the students involved in the interviews:

From Teacher Journal: "8th class"

"After today's lesson (and conclusion) on water quality, I asked the students to self-assess using the rubric. A few of the strugglers were harsh on themselves. One of the boys finished as quickly as he could and put his head down. On asking him what was wrong, he said 'I'm just not good at writing this stuff.' I tried to get him to focus on what he did right and just think of one thing he might do differently next time. Baby steps. Overall though, most of the class seemed to find at least one or two things they could do better. For the top kids it's really just down to complexity and the fact that they haven't learnt the more advanced theories yet. I don't really expect Year 6s to be telling me about Nitrate ions separating in water, so at this stage I need to be cognisant that their Science Theory parts will be 'Year 6' level."

The boy mentioned in the first part of the excerpt was upset with himself and withdrew after a self-assessment task for which he perceived inadequacy. Part of this was exacerbated by his own emotional state of mind on that particular day. This said, the act of self-assessing could be linked to reinforcement of feelings of failure. The refocussing of the student to the positives (the 'stars' from 'star-and-step') was deemed an important part of the process of supporting him through his difficult time. He also needed some management of his expectations. The purpose of asking him to focus on just one area at a time was to make him see some improvement even if his overall writing ability was not enough to see a jump on the assessment scale.

In the second part of the excerpt, a limitation is recognised in the amount of improvement possible by the students in the study. According to the level descriptors in the Conclusion assessment continua used by the school, achievement past level 6 is really dominated by the quality of the explanation of the scientific theory behind the experiment. Many of the top students in the class had all the hallmarks of excellent conclusions, but their lack of experience with more advanced concepts limited the overall complexity of the arguments.

The final journal entry was written after the 'finish' of the intervention strategies in this research:

From Teacher Journal: "Overall reflection"

"I just didn't get enough chances to get the students to do the self-assessment that I would have liked. All of this takes time, and I've been also trying to get content, and other skills like graphing, etc done. I personally think these students are going to need years of practise before they get good at reading a rubric and passing judgement on their own work or that of another. Plus, having the maturity to do this without letting the fact that you are mates with the person you are assessing and having that cloud your judgement.....they need time on that front too. But the acronym – wow – what a difference. You just need to look at the first conclusion they wrote and compare it to their last one. Even if it's not that much better theory-wise, it at least reads like a conclusion and actually does what it's meant to.

For the top performing kids, it was really easy to sculpt their writing into something structured and 'on-the-money' but for my lower ability kids, it still takes some doing. The biggest difference seems to be in their ability to link science theory to what happened in the experiment – there is just a 'disconnect' there that I can't seem to get my head around. Perhaps the concepts are too abstract for them to link the concrete data they are given/collecting? Still, in terms of the other parts of the writing, they are better than they were when they started.

My observations of interactions in class, before assessments and writing activities is that the F.R.E.S.H. acronym is burnt into their memories. They can recite it verbatim and know the meaning of each part– they know what they need to write, and they all seem like they recognise it as the best strategy for structuring their writing."

The F.R.E.S.H. acronym is cited here as the prime-mover in terms of the improvement of the students' writing of scientific conclusions. The peer/self-assessment was considered a good practice. It created experiences for the students to reinforce the importance of the structure of the scaffold and presented them with real examples of good (and bad) of conclusions with which they could improve their own writing. The issues of time, maturity and contact with advanced scientific theory are summarised as the main problems encountered during the intervention process. Some mention is also made of a different rate of improvement for students of different ability/entry levels at the beginning of the intervention process.

4.5 Quantitative Results

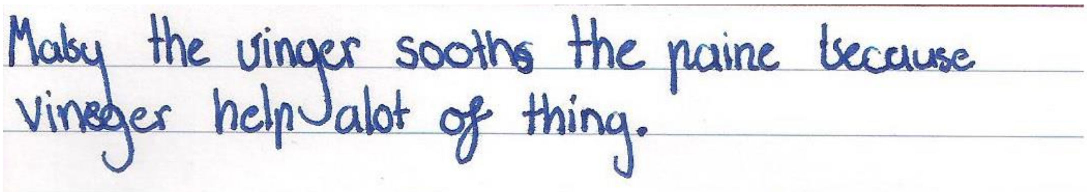
The conclusions written by the students in their journals and the final assessment piece were evaluated by the researcher and given a numerical score based on the Conclusion Writing Continuum (Table 3.6) based on their quality. As all pieces were evaluated by the same person, it is assumed that the same qualitative judgements were used to assign the levels allow for valid comparison and analysis. The selected conclusions represent those written by the majority of the class (on occasion students were absent) under assessment conditions with no help or guidance from the researcher or their peers.

To maintain authenticity and to allow for a link between the quantitative and qualitative results of this research, students' pseudonyms have been preserved with the data.

The preparatory information/data given to the students for each task are given in Appendix C. The 'Jellyfish' and 'Hydrogen' tasks were given to the students with secondary data, whereas the conclusions to the 'Candle' and 'Water Quality' tasks were written after the experiment had been performed in the previous class by the students (primary data). They are presented in chronological order with the 'Jellyfish' task being the diagnostic piece performed before the introduction of the intervention strategies. The 'Candle' task was performed directly after the introduction of the F.R.E.S.H. scaffold and modelling of strong and weak samples. The 'Water Quality' task was performed after the introduction and use of peer/self-feedback using the co-constructed rubric. The 'Hydrogen' task was performed at the end of the intervention process after several opportunities for metacognition and continued use of the other improvement strategies and was the summative assessment used in consequent grading.

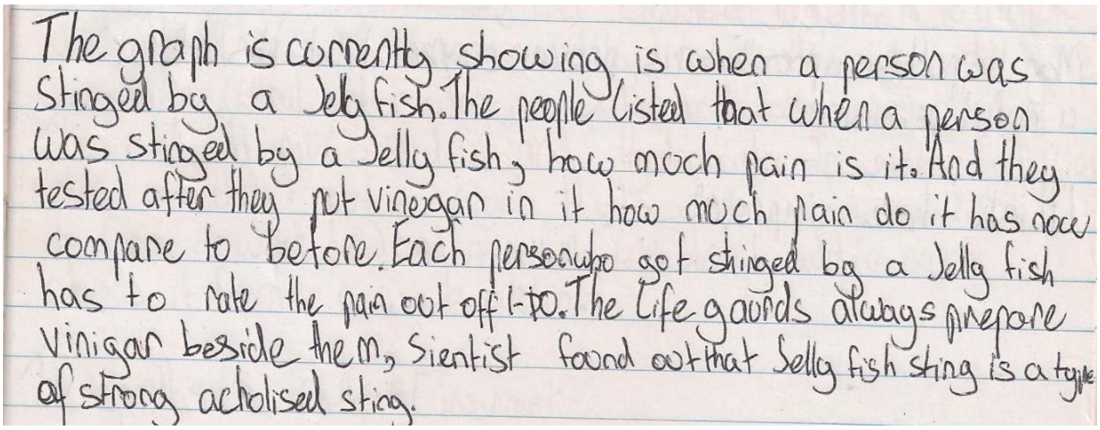
Examples from the different writing tasks and at different levels are provided in Figures 4.7, 4.8, 4.9, 4.10, 4.11 and 4.12. Tasks of the same level are also provided for comparison.

Figure 4.7 ***Writing Sample for the Jellyfish Task Evaluated as Level 1 (From Ralph's Journal)***



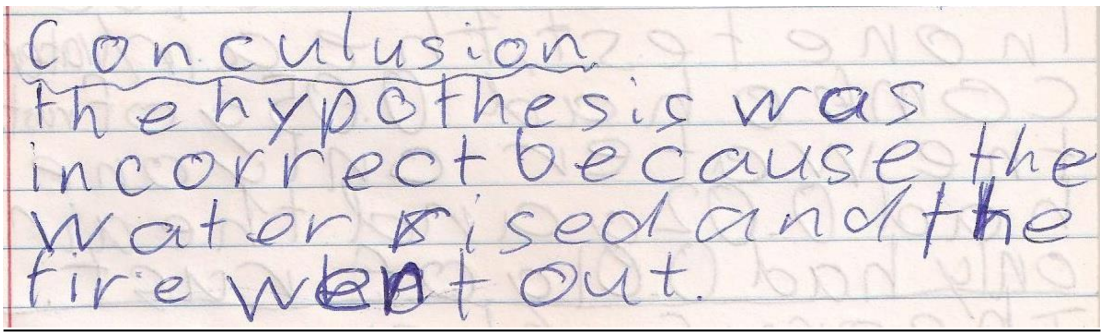
Maby the vinger sooths the paine because
vineger help alot of thing.

Figure 4.8 ***Writing Sample for the Jellyfish Task Evaluated as Level 2 (From Caleb's Journal)***



The graph is corrently showing is when a person was
Stinged by a Jelyfish. The people Listed that when a person
was stinged by a Jelly fish, how moch pain is it. And they
tested after they put vinegar in it how much pain do it has now
compare to before. Each personwho got stinged by a Jelly fish
has to rate the pain out off 1-10. The Life gaurds always prepare
Vinigar beside them, Sientist found out that Jelly fish sting is a type
of strong acholised sting.

Figure 4.9 ***Writing Sample for the Candle Task Evaluated as Level 3 (From Matt's Journal)***



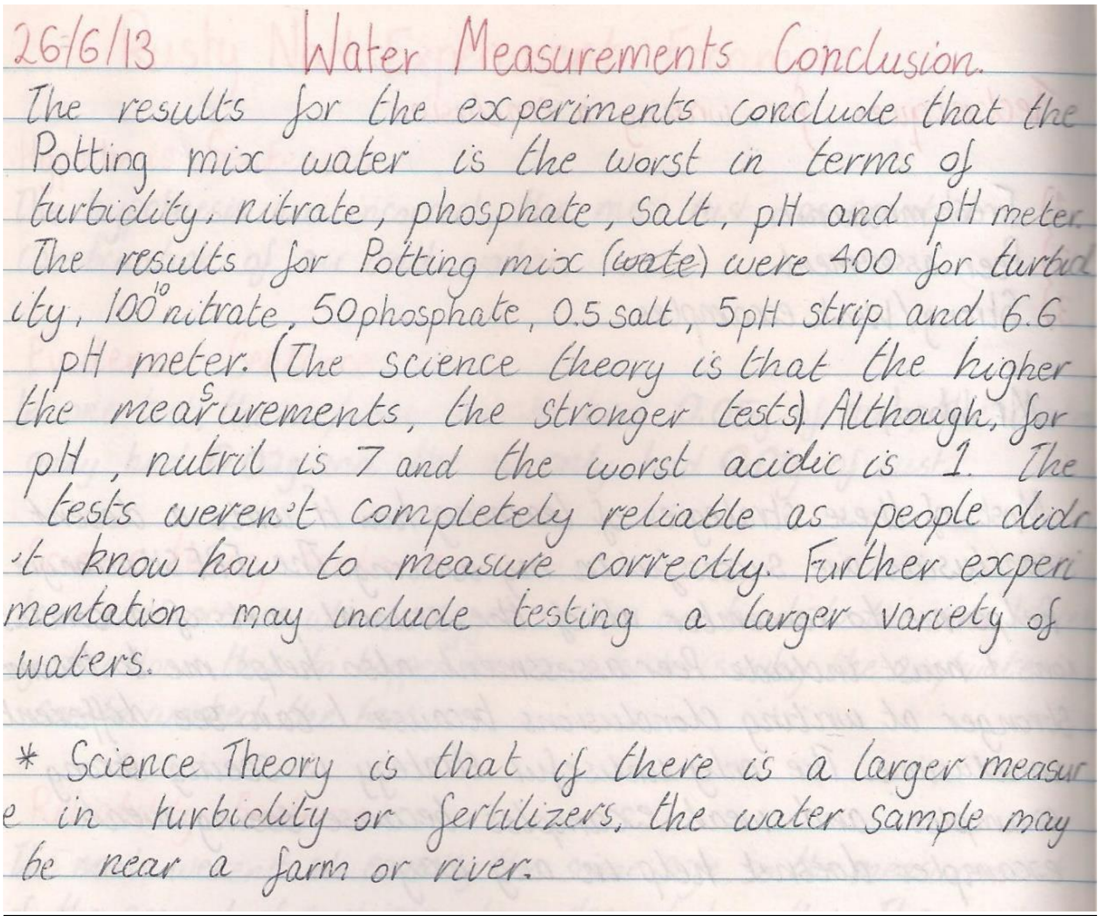
Conculusion
The hypothesis was
incorrect because the
water rised and the
fire went out.

Some limitations of the evaluations are apparent. If you compare the writing of Matt and Beth in Figures 4.9 and 4.11, it is obvious that Beth's standard of writing is more advanced. However, based on the level descriptors in the Conclusion Writing.

Continuum (Table 3.6), they received the same level of achievement. In this case, Beth forgot to support her generalisation with data and was penalised. This said, for the most part the writing quality of similarly evaluated conclusion seems to be the same (for example, Arthur and Allison show similar features in Figures 4.10 and 4.12). As cohesiveness and efficiency of the writing were not assessed, it is hard to make claims regarding these facets of the writing.

Four common writing tasks were selected for comparison. The achievement scores each student received, based on the Conclusion Writing Continuum (Table 3.6), for each task is displayed in Table 4.1. The data clearly shows a shift in the overall quality of the writing for the class based on the mean level of achievement over the course of the intervention process. A striking gain in student achievement (almost double) is seen after the introduction of the F.R.E.S.H. scaffold and modelling of the writing. Interestingly, subsequent evaluations show a tailing off of the overall achievement of the class. It is hard to attribute this effect to the introduction of the peer/self-assessment and opportunities for metacognition. The instrument was used on four different occasions to measure the student progress against the Conclusion Writing Continuum (Table 3.6). The changing nature of the science curriculum taught during the intervention period dictated that the instrument had to be changed to match the subject matter. Although the four tasks are essentially measuring the same skill (conclusion writing), it seems more likely that variations in the difficulty of the subject matter and the nature of each task have slightly affected the criterion validity of the instrument. The researcher does not deem the variation from task to task to be pronounced enough to place the findings in any doubt and it is only mentioned in the interests of transparency.

Figure 4.10 ***Writing Sample for the Water Quality Task Evaluated as Level 7 (From Arthur's Journal)***

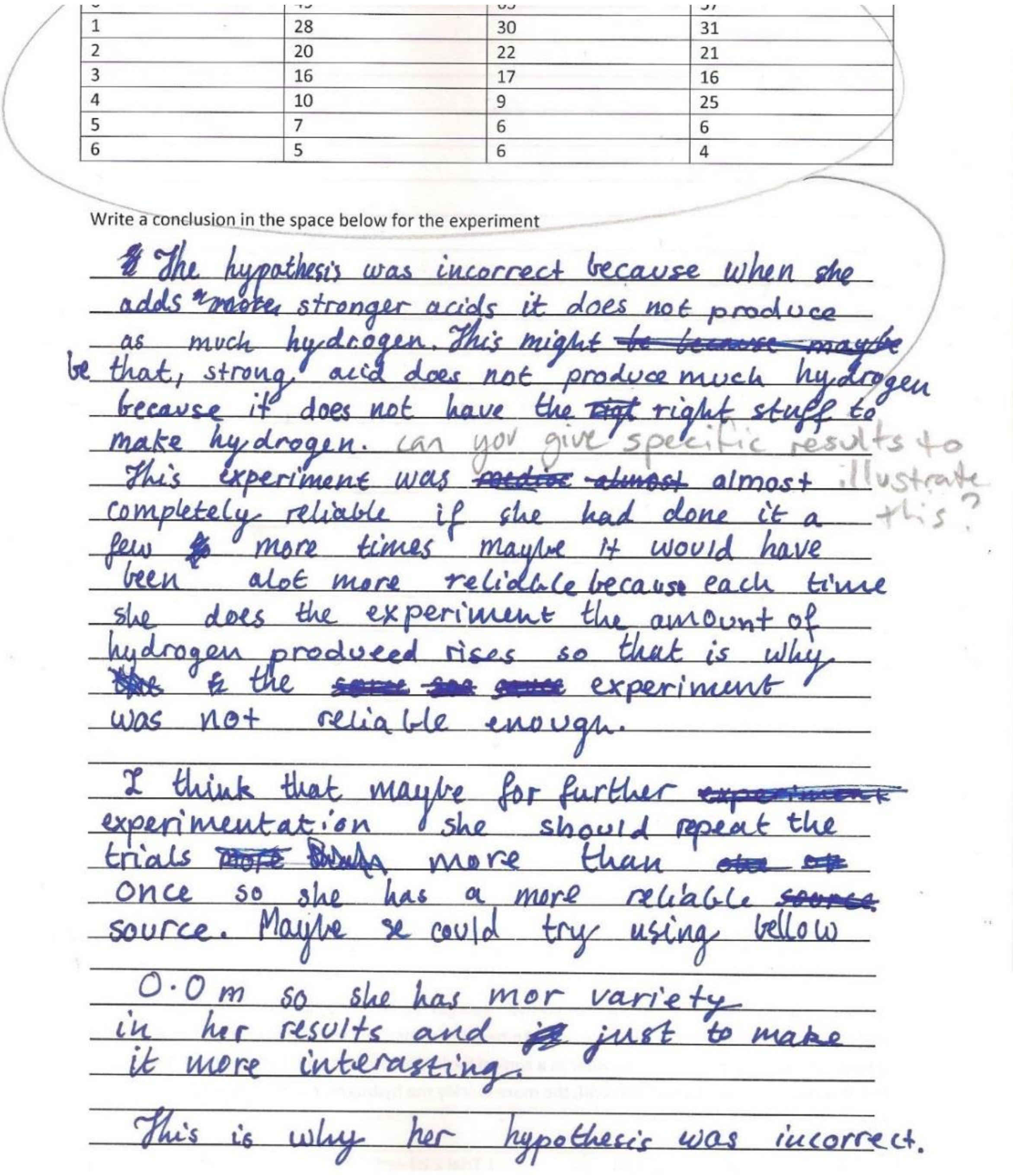


26/6/13 Water Measurements Conclusion.

The results for the experiments conclude that the Potting mix water is the worst in terms of turbidity, nitrate, phosphate, salt, pH and pH meter. The results for Potting mix (wate) were 400 for turbidity, 100 10 nitrate, 50phosphate, 0.5 salt, 5pH strip and 6.6 pH meter. (The science theory is that the higher the mearurements, the stronger tests) Although, for pH, nutril is 7 and the worst acidic is 1. The tests weren't completely reliable as people didn't know how to measure correctly. Further experimentation may include testing a larger variety of waters.

* Science Theory is that if there is a larger measure in turbidity or fertilizers, the water sample may be near a farm or river.

Figure 4.11 ***Writing Sample for the Hydrogen Task Evaluated as Level 3 (Beth)***



1	28	30	31
2	20	22	21
3	16	17	16
4	10	9	25
5	7	6	6
6	5	6	4

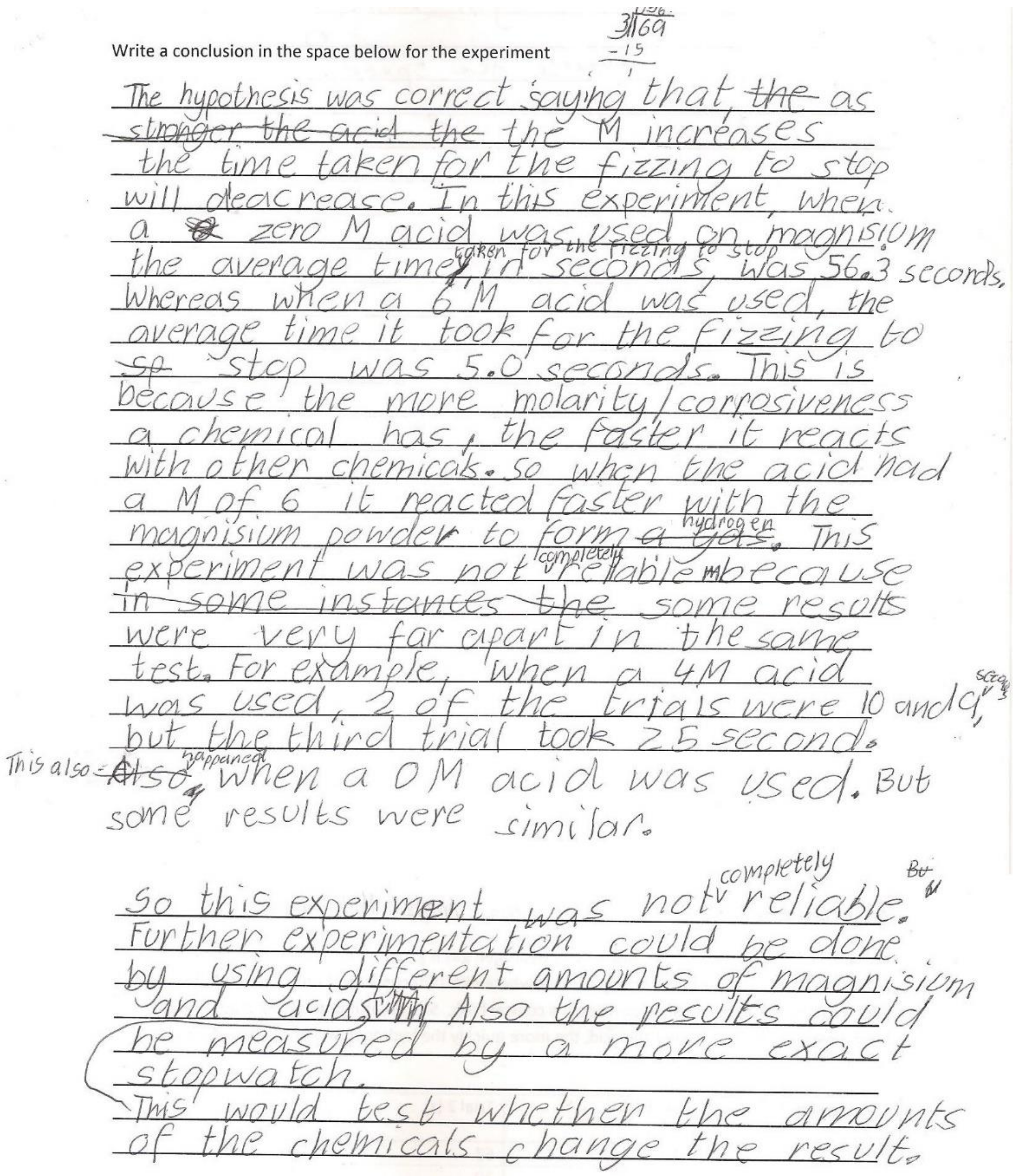
Write a conclusion in the space below for the experiment

The hypothesis was incorrect because when she adds more stronger acids it does not produce as much hydrogen. This might ~~be because maybe~~ be that, strong acid does not produce much hydrogen because it does not have the ~~rigt~~ right stuff to make hydrogen. *can you give specific results to illustrate this?* This experiment was ~~almost~~ almost completely reliable if she had done it a few more times maybe it would have been alot more reliable because each time she does the experiment the amount of hydrogen produced rises so that is why the experiment was not reliable enough.

I think that maybe for further ~~experiment~~ experimentation she should repeat the trials more than once so she has a more reliable ~~source~~ source. Maybe se could try using bellow 0.0 m so she has mor variety in her results and just to make it more interasting.

This is why her hypothesis was incorrect.

Figure 4.12 ***Writing Sample for the Hydrogen Task Evaluated as Level 7 (Allison)***



Write a conclusion in the space below for the experiment

The hypothesis was correct saying that as the M increases the time taken for the fizzing to stop will deacrease. In this experiment, when a zero M acid was used on magnisium the average time taken for the fizzing to stop in seconds was 56.3 seconds. Whereas when a 6 M acid was used, the average time it took for the fizzing to stop was 5.0 seconds. This is because the more molarity/corrosiveness a chemical has, the faster it reacts with other chemicals. So when the acid had a M of 6 it reacted faster with the magnisium powder to form hydrogen. This experiment was not completely reliable because some results were very far apart in the same test. For example, when a 4M acid was used, 2 of the trials were 10 and 9 secs, but the third trial took 25 seconds. This also happened when a 0M acid was used. But some results were similar.

So this experiment was not completely reliable. Further experimentation could be done by using different amounts of magnisium and acid. Also the results could be measured by a more exact stopwatch. This would test whether the amounts of the chemicals change the result.

Tables 4.2 and 4.3 separate the data (achievement scores for the four selected writing samples) by gender and the mean scores are calculated for comparison.

It is worth noting that, although the female student show quite a substantial difference in their levels of achievement when compared to the males, they also start with a higher entry point with a mean of 3.1 as compared to 2.1. Also, the boys actually outperformed the girls (slightly) in the 'water quality' task. The practical nature of the water quality task may have appealed to the male students more than girls (many of whom deemed it 'yucky').

Table 4.1 ***Student Achievement Scores on Conclusion Writing Tasks.***

Student	**'Jelly Fish' Task**	**'Candle' Task**	**'Water Quality' Task**	**'Hydrogen' Task**
Beth	3	5	4	3
Lucy	4	7	5	4
Matt	2	3	2	3
Peter	3	6	5	4
Karen	1	4	2	4
Arthur	3	7	7	7
Tina	4	6	7	6
Steven	2	5	4	3
Rowena	3		3	6
Abigail	4	5		6
Allison	4	8	7	7
David	2	3	4	3
Ralph	1	3		2
Michelle	3	5	3	3
Bradley	1	3	3	4
Nathan	3	7	7	5
Eleanor	2	5	2	2
Jenny	3	4	7	5
Caleb	2	3		3
Mean score	**2.6**	**4.9**	**4.5**	**4.2**

The student levels were sorted by their performance on the initial task 'Jellyfish' and then by their mean score across all four tasks in order to provide some semblance of an entry rank into the intervention process. The range of students' scores was calculated by subtracting their lowest level of achievement from their highest. Table 4.4 shows the sorted achievement scores and subsequent calculations of mean and range.

Table 4.2 ***Student Achievement Scores on Conclusion Writing Tasks Sorted by Gender (Male).***

Student	**'Jelly Fish' Task**	**'Candle' Task**	**'Water Quality' Task**	**'Hydrogen' Task**
Matt	2	3	2	3
Peter	3	6	5	4
Arthur	3	7	7	7
Steven	2	5	4	3
David	2	3	4	3
Ralph	1	3		2
Bradley	1	3	3	4
Nathan	3	7	7	5
Caleb	2	3		3
Mean score	**2.1**	**4.4**	**4.6**	**3.8**

Table 4.3 ***Student Achievement Scores on Conclusion Writing Tasks Sorted by Gender (Female).***

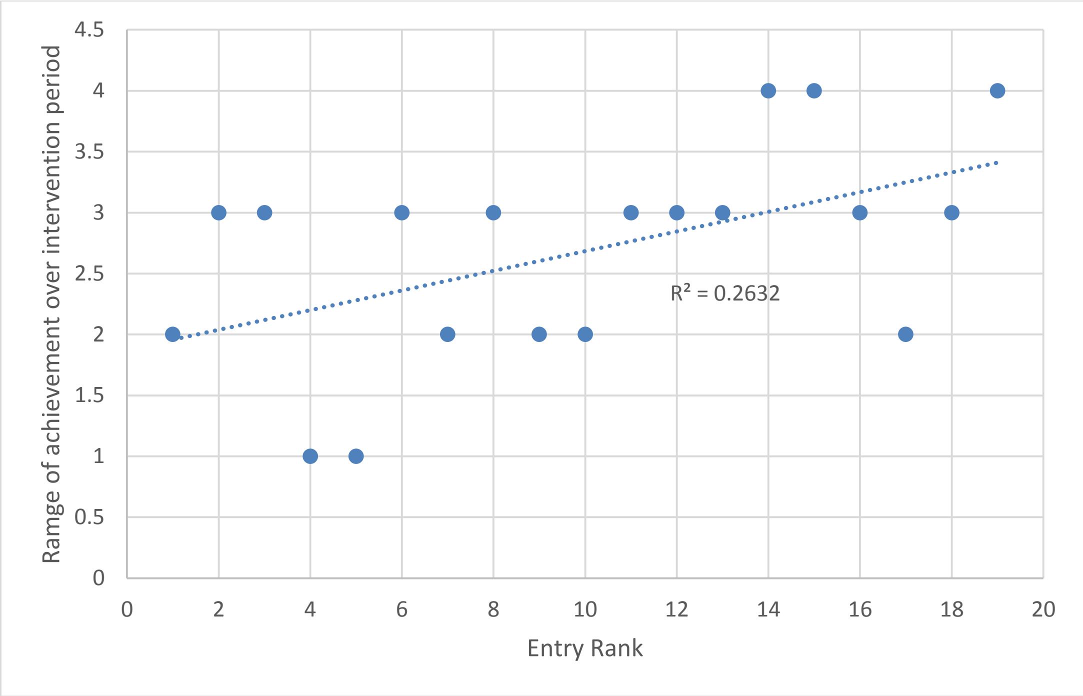
Student	**'Jelly Fish' Task**	**'Candle' Task**	**'Water Quality' Task**	**'Hydrogen' Task**
Beth	3	5	4	3
Lucy	4	7	5	4
Karen	1	4	2	4
Tina	4	6	7	6
Rowena	3		3	6
Abigail	4	5		6
Allison	4	8	7	7
Michelle	3	5	3	3
Eleanor	2	5	2	2
Jenny	3	4	7	5
Mean score	**3.1**	**5.4**	**4.4**	**4.6**

Table 4.4 ***Student Achievement Scores on Conclusion Writing Tasks Sorted by 'Entry Rank'***

Student	**'Jelly Fish' Task**	**'Candle' Task**	**'Water Quality' Task**	**'Hydrogen' Task**	**Mean (for student)**	**Range**
Ralph	1	3		2	2	2
Karen	1	4	2	4	2.75	3
Bradley	1	3	3	4	2.75	3
Matt	2	3	2	3	2.5	1
Caleb	2	3		3	2.67	1
Eleanor	2	5	2	2	2.75	3
David	2	3	4	3	3	2
Steven	2	5	4	3	3.5	3
Michelle	3	5	3	3	3.5	2
Beth	3	5	4	3	3.75	2
Rowena	3		3	6	4	3
Peter	3	6	5	4	4.5	3
Jenny	3	4	7	5	4.75	3
Nathan	3	7	7	5	5.5	4
Arthur	3	7	7	7	6	4
Lucy	4	7	5	4	5	3
Abigail	4	5		6	5	2
Tina	4	6	7	6	5.75	3
Allison	4	8	7	7	6.5	4
Mean	**2.6**	**4.9**	**4.5**	**4.2**	**4.0**	**2.7**

In order to determine whether the entry rank correlated with the amount of improvement, the performance on the initial task (sorted by level, and then by mean performance overall) was graphed against the range of achievement for each student using Microsoft Excel (2013). The graph is presented in Figure 4.13.

Figure 4.13 ***Graph of Entry Rank vs. Range of Achievement***



Microsoft (2013)

The Pearson product moment correlation coefficient was calculated using by Microsoft Excel and can be found taking a square root of the r-squared value on Figure 4.13. The following quote provides a description of the Pearson function utilised in the program:

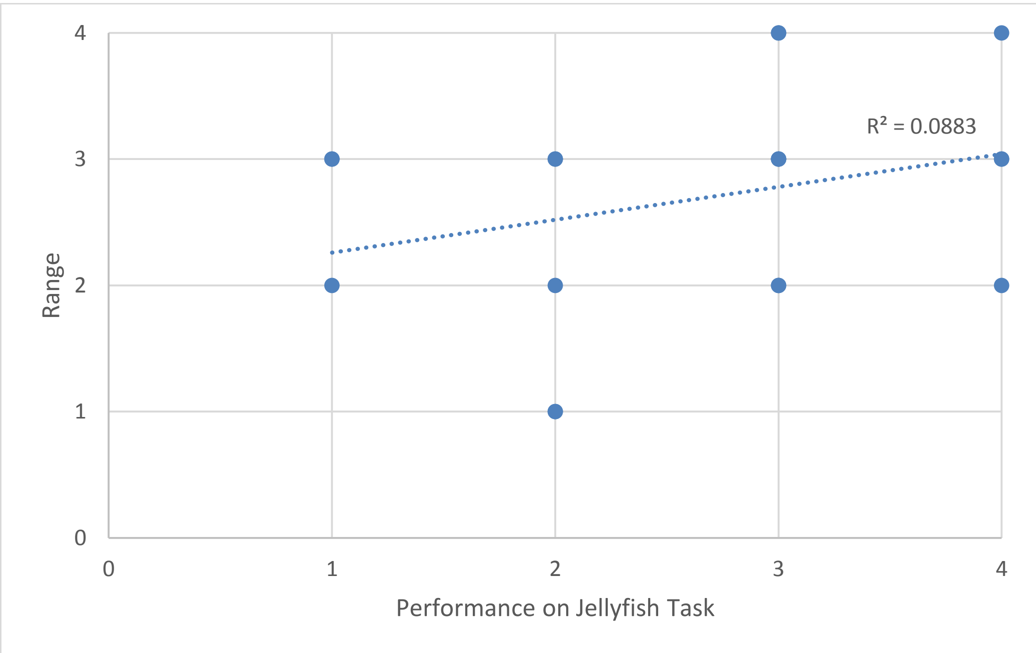
> "The r-squared value can be interpreted as the proportion of the variance in y attributable to the variance in x…
>
> The equation for the Pearson product moment correlation coefficient, r, is:
>
> $$r = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sqrt{\sum (x-\bar{x})^2 \sum (y-\bar{y})^2}}$$
>
> where $\bar{x}$ and $\bar{y}$ are the sample means AVERAGE(known_x's) and AVERAGE (known_y's). RSQ returns R^2, which is the square of this correlation coefficient." (Microsoft, 2013)

The Statistics How To website (2015, accessed 12 March 2015) gives a relatively easy way to interpret the Pearson product moment correlation coefficient:

> "The results will be between -1 and 1. You will very rarely see 0, -1 or 1. You'll get a number somewhere in between those values. The closer the value of r gets to zero, the greater the variation the data points are around the line of best fit.
>
> High correlation: .5 to 1.0 or -0.5 to -1.0
>
> Medium correlation: .3 to .5 or -0.3 to -.5
>
> Low correlation: .1 to .3 or -0.1 to -0.3"
>
> (Statistics How To, 2015, accessed 12 March 2015)

With a Pearson product moment correlation coefficient of 0.5, it would seem that there is a medium to high correlation between the entry point into the intervention process in terms of writing ability (when compared to the rest of the class) and the amount of improvement that a student may achieve.

Figure 4.14 ***Performance on Jellyfish Task vs. Range of Achievement***



Microsoft (2013)

A second correlation, comparing achievement on the initial evaluation of writing ability (the Jellyfish task) and the range of achievement, was performed and is presented in Figure 4.14. It should be noted that the Pearson product moment correlation coefficient for this second correlation is 0.3, which indicates a low to medium correlation between the performance on the initial task and the range of the performance on the other tasks.

4.6 Summary

Chapter four was divided into four major sections, one for each instrument used to measure changes in the students writing. The first instrument considered was the interviews with selected students before and after the introduction of the writing intervention strategies in Section 4.2. The interviews indicated that the students perceived an improvement in their writing as a result of the strategies. Section 4.3 involved an analysis of the students' journals and their reflections on their writing and the intervention strategies. The third instrument was that of the researcher's own journal and was presented as a narrative of the improvement journey in Section 4.4.

The final instrument was the quantitative analysis of the student's writing taken over the course of the research project. Section 4.5 gave the numerical results indicated a substantial gain in the achievement level of the students overall based on the Conclusion Writing Continuum (Table 3.6). It was found that there was a moderate correlation between the amount of improvement and the initial writing ability.

The next chapter finalises the thesis by providing a conclusion and discusses the implications of the research.

CHAPTER 5: CONCLUSION AND IMPLICATIONS

5.1 Introduction

The final chapter of the thesis provides the reader with the conclusions that have been drawn from the research and their implications. It is organised into three main parts. The first section deals with the research questions and the extent to which they have been answered. Each question is considered in turn and evidence is used to support claims made by the researcher. The second section discusses how this research has contributed to the knowledge of educators and the implications for the research participants, the participating school and the researcher himself. In the final section, issues with the research project are described and recommendations for further research are made.

5.2 Answering the Research Questions

The research questions are rephrased as the title for each part and are considered in turn.

5.2.1 How Students' View of Themselves as Scientific Writers Change?

The students see themselves as improved scientific writers. The main evidence for this comes from the interviews conducted with the students after the implementation of the intervention strategies, the reflections made by the students in their journals and from the researcher's observations of student reactions in the teacher journal.

All of the students interviewed stated that they felt that they had improved as a science writers. They could describe what makes a good conclusion by listing the structural components of the F.R.E.S.H. acronym and recognised how different their writing was from before the intervention strategies. As the students considered these differences, they could all state that their writing was of a higher standard. As can be seen in the quantitative data, all of the students made progress in some form or another. It would be hard to imagine that any of them would think they had not made progress.

The students initially regarded scientific writers as people who had large vocabularies and wrote reams of work. Many of the themes that came through in the initial interviews involved problems students that had with the vocabulary of science and the idea that quantity meant quality. The final interviews saw a shift in students' attitudes to the writing. They began to see that the content of their writing and how they structured it logically and coherently were of more importance.

Most of the reflections made by the students late in the research project indicated that they saw themselves as better writers. Jenny's remark "I think I have improved in making conclusions [sic]" typifies the sort of thinking the students wrote in their journals. Some of the reflections connected the amount of improvement with the level of use of the strategies for constructing a writing piece such as the F.R.E.S.H. scaffold. There were no comments that indicated any lack of development in the students' opinion of themselves as scientific writers.

In the teacher journal, the students' ability to describe the structure of a good conclusion is described as being "burnt into their heads", referring to the repetitive use of the F.R.E.S.H. scaffold. From the first introduction of the scaffold, the researcher noticed a change in the way in which the students approached their writing. They were able to write with more purpose and a clearer understanding of the communication objectives that a conclusion had to provide the reader.

The researcher acknowledges that, in terms of 'scientific writing', the conclusions written by the students do not encompass all of the possible modes that scientific discourse might take. A conclusion (a generalisation supported by evidence and theory) is deemed by this researcher to be a very central tenet of the idea of scientific writing. Therefore improvement in the writing of a conclusion is an important part of improving the way in which the students communicate scientifically.

Rice (1998, p. 271) made note of a change in student perceptions of their writing after using a scheme of self-assessment to evaluate their work:

> "Students often say that they began the course thinking that they wrote rather well, but that my comments quickly disabused them of that opinion. From their

> vantage point at the end of the semester, they are usually surprised at how unsuccessful their first papers were, how much their writing improved over the course of 13-14 weeks, and how important revision is to the writing process."

After keeping a written record of their work for approximately 10-11 weeks in their journals, the students in this research project were able to reflect on their early pieces of writing and make comparisons between their later work and earlier work. This has definitely changed their perceptions of themselves as 'science writers' in a positive way.

5.2.2 How Will Students' Self-efficacy Towards Scientific Writing Change?

The majority of the students felt more confident and better able to write in science. They all experienced success and did not view scientific writing as negatively as they may have done. In the interviews with students before the intervention strategies, the lower achieving students typically had low expectations of their ability to write. For example, in his first interview, David said "it's a low standard or a medium standard because I'm not putting in a lot of effort or all the information in." After the intervention process the students' view of themselves improved. In his last interview, David referred to his successes saying, "It makes me feel a bit more like… 'in science', than before".

In terms of exerting control over their learning, the students believed that they were able to make changes in order to achieve better results. They recognised strategies that help them to write more structured conclusions. The metacognitive excerpts from their journals are the main sources of evidence for this. Comments like "Because I didn't really use the FRESH acronym so I didn't improve much" indicate that students recognised which behaviours are likely to lead to success and could identify what needs to change about their current behaviours.

One student (Ralph) indicated early in the project through the journal entries that he felt as though he wasn't very good at science. This was in response to a self-evaluation of the first pieces of writing. After subsequent improvements in his writing over the intervention period, his reflections did not have the same negative views. This could be due to a change in his belief in his abilities.

Bandura (1993, p. 125) described how the nature of feedback can influence self-efficacy saying, "Performance feedback that focuses on achieved progress underscores personal capabilities. Feedback that focuses on shortfalls highlights personal deficiencies." The nature of the journal as a record of students' progress has allowed students to focus on 'where they have come from' rather than pointing out how much further they 'have to go'. Also, the peer/self-assessment rubric allows students to see how much they had achieved rather than solely focussing on what their writing was lacking.

There was possibly some contradictory evidence to the overall claim of improved self-efficacy. An example of low self-worth was exhibited by a student presented in the teacher journal. A student declared after a self-assessment, "I'm just not good at writing this stuff". This was towards the end of the intervention process when multiple peer/self-assessments of their writing had been conducted. It is possible that this student, who had not achieved very highly when compared to his peers, was given a constant reminder of his position in the class and that this reinforced ideas of limited achievement. This could potentially have had a negative effect on his confidence in himself when writing. It must also be considered that this may have been an isolated incident for which he was reacting to just that particular writing piece. As the outburst was not followed up with further questioning, these assumptions remain speculation.

5.2.3 How Will Students' Performance in Scientific Writing Change?

The students' writing performance has improved dramatically with the introduction of the F.R.E.S.H. scaffold. The quantitative results showed that the mean level of achievement on the Conclusion Writing Continuum (Table 3.6) jumped from an initial 2.6 (the Jellyfish task) to 4.9 on the very next evaluation (the Candle task). The main cause of this is attributed to the introduction of the F.R.E.S.H. scaffold and modelling of strong and weak samples. The achievement remained high (mean above a 4) for the remainder of the research project.

There was some variation in the mean for subsequent evaluations of student performance in conclusion writing (the Water Quality and Hydrogen tasks), which was most likely due to the differences in the tasks themselves. Criterion validity may have been affected by the familiarity with the scientific concepts behind the experiments, the use of first- and second-hand data and the level of engagement by the students.

The effect of peer/self-assessment did not seem to have made as dramatic a change to the achievement on the Conclusion Writing Continuum. Given that the scaffold and modelling was presented to the students first, it is natural that it would be most likely to show the largest change in their achievement scores. It is thought by the researcher that the peer/self-assessment helped to reinforce and consolidate the gains provided by the scaffolding and modelling.

The metacognitive journaling is in the same league as the peer/self-assessment in terms of its 'effect size' on the students' ability to write a scientific conclusion. The quantitative results did not show a distinct gain in student performance as a result of the implementation of the journaling. Again, it is thought that the effects of the strategy were combined with the gains made by the scaffold and modelling strategy in the form of reinforcement.

The research conducted seems to confirm the findings of many other educational researchers. The strategies that were employed to improve the students' writing succeeded in their intent. Similarities with other research are considered below.

Early success was enjoyed after the introduction of F.R.E.S.H. scaffold. Immediate gains in the students' level of achievement on the Conclusion Writing Continuum were observed. Rutherford (2007) also noted rapid gains after the introduction of writing guidelines and attributed this to unfamiliarity with the requirements of scientific writing. The majority of the Year 6 students in this study had previously studied science (formally) for one year prior to the project at the research school. It seems likely that the unfamiliarity with scientific writing requirement was the main impediment to their achievement.

The study performed by Hand, Wallace and Yang (2004) using their Science Writing Heuristic (SWH) (Keys, Hand, Prain & Collins, 1999) showed marked improvement in scientific writing for the students that used it to structure their writing. The F.R.E.S.H. scaffold has similar facets to the SWH in that it provides a writing template with descriptions for each segment.

To model the F.R.E.S.H. scaffold students were provided with writing samples of varying complexity and quality. This was seen as an amalgam of the work of LaConte and Berry (2006) and Brandt (1971) in which writing was not only modelled as a strong sample, but also as a weak sample which had flaws introduced. This was done as an introduction to the critiquing techniques that the students would need to employ during peer/self-assessment. Similarly, the results of this study have shown that the provision of such writing samples and the modelling of correct writing have resulted in improved scientific writing for students.

Peer/self-assessment had mixed success in this research project. Initial use of peer assessment was found lacking due to the students' inexperience and maturity. Rice (1998) had found peer assessment useful at the university level where it is thought (by this researcher) that the 'need-to-please' exhibited by the young students in this study would have been minimal. The introduction of a rubric to guide their assessments was found to be very helpful as in the research conducted by Butler and Nesbit (2008), who used a co-constructed rubric to help students to self-assess. Once students utilised the rubric to evaluate each other's (or their own) writing, the feedback became more focussed and the next steps for improvement became readily apparent.

The male students seemed to have improved more than the female students. When considered by gender, the initial difference between the mean for the Jellyfish and Candle tasks was 2.3 for both the males and females. The difference between the initial task and subsequent tasks was more varied between the sexes, however. The difference between the means of the Jellyfish task and the Water Quality task was 2.5 for the males and 1.3 for the females. The difference between the means of the Jellyfish task and the Hydrogen task was 1.7 for the males and 1.5 for the females. When considering these results, it should be noted that the male students started from a lower mean level than the females, and that the Conclusion Writing Continuum was by no means a linear scale of improvement. It could simply be 'easier' to make gains on the Conclusion Writing Continuum from a lower starting position. Consequently it is difficult to make any claims that the scientific writing improvement strategies employed in this research project were differentially effective for males and females.

Along these same lines, brief analysis of entry rank and improvement was conducted and it was found that there was a medium to strong correlation (R value of 0.5) between the initial level of ability (compared to the rest of the class) and the amount of improvement (in terms of range) achieved by a student. This was supported by observations made by the researcher in the teacher journal, stating that, "For the top kids, it was really easy to sculpt their writing into something structured and 'on-the-money' but for my low kids, it still takes some doing". When the initial level of ability is directly compared to the range for each student the correlation is weak-medium (R-squared value of 0.3). The intervention strategies seem to favour students that already have substantial writing skills, channelling that ability into effective scientific communication. For the weaker students, the strategies do not seem as effective.

This research has extended knowledge in the area of improving scientific writing on two main fronts. Firstly, a new way of scaffolding and assessing student writing (the F.R.E.S.H. scaffold) has been shown to be highly successful in building student confidence in their own abilities as scientific writers, as well as changing the substance of their writing to a more coherent and logical format. Secondly, the research has shown that delivery of a multi-faceted intervention process with overlapping strategies can successfully change the outcomes for student learning in the area of scientific writing. Nineteen students and one teacher have vastly changed the way they approach learning and teaching about scientific conclusion writing. With distribution of the results of the research, other teachers and students may also change their approach as well.

5.2.4 What Problems Arise From the Implementation of the Intervention Strategies?

The main source of information used to answer this research question was the teacher journal kept by the researcher over the course of the implementation of the intervention strategies. One of the first noticeable problems was the amount of time taken to train the students to effectively use peer/self-assessment. The researcher found that the students had very little skill in identifying how to give meaningful feedback. The initial use of peer assessment was observed to be superficial and dominated by the students desire to make each other feel good with congratulatory remarks which were observed in the students' journals. This form of feedback was deemed ineffectual by the researcher. To combat this, the researcher's response was to develop a structured way of giving feedback by co-constructing a marking rubric to serve two purposes. It reinforced the F.R.E.S.H. scaffold by using the mnemonic as its main criteria for evaluation. It also allowed the students to rate each other as below, at or above the expectations of the class for writing a conclusion. In retrospect, the researcher regrets the choice of words used to rate each criterion. It is felt that students might feel ashamed at being described as 'below expectations' and that this might contribute to negative self-efficacy surrounding writing scientifically. The rubric also allowed space for general comments to be made utilising a star (what was done well) and step (what should the student focus on next time) method. The comments provided here were generally related to the progress made in the rubric or some form of congratulatory remarks.

The construction of the rubric took time both in and out of class, which meant that less time was spent on other content. The same could also be said of the amount of time allowed for students to write in their journals and for analysing the writing samples and the F.R.E.S.H. scaffold. The researcher was cognisant of this and described his dilemma in the teacher journal. There was a constant question that played on the researcher's mind. Is the amount of time that is spent on improving writing causing a decline in other areas of the Science curriculum for the participating students? A detailed study of students' progress in other areas was beyond the scope of the research project and so no claims can be made quantitatively. This said, no decline in achievement scores on other assessments was noticed by the researcher that would signify deterioration of their scientific knowledge, skills and understandings due to extra time spent on the writing intervention tasks.

It was noted in the teacher journal by the researcher that feedback from the students tended to be better from the high-achieving students rather than the low-achieving students. The researcher felt that this inequality undermined the effectiveness of the peer assessment as an improvement strategy. Although the rubric did a lot to equalise this, the comments provided by students were often more detailed when given by a high-achiever. Another source of inequality/inaccuracy in the feedback was when it was provided by a 'friend'. It was noticed by some students (and commented on in their journals) that the marks tended to be higher or more lenient depending on the person giving the feedback and his/her relationship to the writer. From a purely logistical point of view, it was deemed easier to swap writing samples with names rather than collect them all and distribute them anonymously. The provision of immediate feedback was considered more important to the researcher than the bias that might ensue from knowledge of the author's identity.

The relative worth of provision of weak writing samples was called into question a number of times. Comments made by the students in the interviews and their journals and observations made by the researcher in the teacher journal indicated that students felt that the weak writing samples were a waste of time. This thinking is probably best summed up by Arthur's excerpt from his journal where he said, "seeing weak examples doesn't help me in any way." Many, but not all, of the students seemed to believe that looking at writing samples with mistakes or poor explanations did not help them to write stronger conclusions; instead it was just 'wasting their time'. The researcher believes that the weak samples helped the students to become better peer assessors by training them to see how a sample could be improved.

It was observed by the researcher in the teacher journal that there seemed to be an upper limit to achievement on the Conclusion Writing Continuum used to evaluate the quality of the students' writing. Indeed, when the results are considered in the four conclusions evaluated over the course of the intervention period, it is noticed that the high achieving students seem to hit a maximum score of level 7 with only one exception. The cause of this is put down to the conceptual understandings that the students have at a Year 6 level. The 'S' in the F.R.E.S.H. scaffold stands for 'Scientific Theory'. Achieving beyond level 7 in the continuum requires a more developed and technical understanding of the science behind the results. For the students to attain a higher level, they must describe more abstract concepts than these provided in the standard Year 6 Science curriculum at the Research School.

By itself, the study in this thesis was not designed to compare the relative worth of the strategies employed to improve the scientific writing of the students. This said, when considering the three intervention strategies employed, this research indicates that the largest effect seems to have been from the introduction of the F.R.E.S.H. scaffold. Hattie's (1992) research indicates that feedback should have had a more 'noticeable' effect on the students' work. So why wasn't it? The researcher believes that, in a way, the introduction of the scaffold to the students after the initial work on the Jellyfish task was an instance of extraordinary self-assessment and self-feedback. The students wrote their diagnostic piece and were immediately confronted with a set of guidelines that told them what a 'good conclusion' looked like. The researcher then immediately modelled what they 'should have' written in a co-constructed piece. The students were confronted by the differences between the diagnostic piece that they had written and the modelled piece. They received feedback on their own work immediately and, maybe more importantly, from their own evaluation of their work. Of all the learning opportunities over the course of the intervention process, the researcher considers the one described above to be the most poignant.

5.3 Implications

The implications of the research are considered below in four sections. The first three consider the major stakeholders in the research, the students, the participating school and the researcher. Finally, consideration is given to future research that might be conducted based on the findings of the thesis.

5.3.1 The Students

The students improved their writing skills in science and learnt ways to structure a conclusion so that it is logical and more coherent. The F.R.E.S.H. acronym was "burnt into their memories" to use the words of the researcher in the teacher journal. They are likely to carry this scaffold with them into the future and use it to sculpt their writing so that it will be of much higher quality. Many of them considered this to be an effective method for remembering the structure of a piece of writing and have created their own mnemonics. For example, in his final interview, David described another mnemonic, ‘C.R.A.B.S’, which he and his peers created to describe the structure for writing the methodology of an experiment.

In terms of providing feedback, the students still require some help to give mature and honest appraisals of each other’s and their own work. A rubric was shown to be immediately effective in improving the quality of this feedback. Students need to be constantly reminded that honest and constructive feedback is required for the feedback to have the most impact on their writing skill. It seemed hard for this age group to be able to separate their friendships from this process.

Many of the students had a boost to their self-efficacy surrounding scientific writing. As a consequence of the successes that they all experienced, they perceived a greater degree of control over their learning of this skill. It would seem that the improvement in this area was neither by the same degree nor was it universal. Some observations were made of students who still had negative perceptions of self-worth in this regards, and this may be due to the fact that those students did not have the same range of achievement on the Curriculum Writing Continuum as some of their peers. Some students only improved by one level, whereas others improved by three or four. This may have had an effect on the students, exacerbated by the fact that the improvements were made better known by the use of peer-assessment.

The students became more reflective of their learning. The use of self-assessment and the provision of journaling opportunities allowed the students to examine the work that they had done. They were able to see which strategies were the best for them to improve their work. Most of them have selected the F.R.E.S.H. scaffold as being most effective, whilst some believed that asking for feedback was better. Regardless of their choice, it is important that they had engaged with their learning meta-cognitively and had become better learners for doing so.

5.3.2 The School

The research school benefited from the research as a class of its students now has better skills when writing in science. As these students carry on through the school, it is hoped that they will share their success with others as they spread throughout other classes in the years that follow. The skills that have been learnt in the Science classroom can no doubt be applied to other areas of their schooling with minor modification.

The researcher continues teaching at the school and intends to share his findings with other teachers. It is hoped that the success with this class will be evaluated and possibly embedded into the curriculum. The implications of this would be a widespread increase of the scientific writing ability across the school.

5.3.3 The Researcher

As a participant researcher, the author has learnt a great deal from the research. Initially the researcher was dubious about the effectiveness of using journals for instructional purposes. From a teaching standpoint, the journals have been shown to provide an excellent source of reflexive teaching practice. The journals chronicle the students' learning process. From the initial diagnostic piece of writing to the summative assessment at the end of the intervention process, the journal provided the researcher with insights into the students' thinking and development of their writing skills. With regular collection and consideration of their journals, the researcher was able to see how the students were progressing and make changes to the way in which the intervention strategies unfolded. The development of the Peer/Self-assessment rubric was a direct response to the lack of ability on the students' part to give meaningful feedback. This was clear after the first review of the journals and changes were made immediately.

The researcher has also been made more cognisant of the importance of self-efficacy in the classroom. Experiences such as the one recorded in the teacher journal with a boy saying "I'm just not good at writing this stuff", have caused the researcher to more carefully evaluate the ways in which students receive feedback. Mindful of the range of improvements made by the students during the research project, the researcher made changes to the way suggestions for writing development were delivered to the students. Only one or two improvements were suggested at a time, so as not to overwhelm students if their writing needed substantial work.

Scaffolding and strong/weak writing samples proved to be extremely effective at improving the students' writing of conclusions. This caused the researcher to develop other areas of scientific writing using these strategies with other classes and year levels. Hypothesis, methodology and data interpretation writing scaffolds were developed with the students to help with these areas. This represents a shift in the types of learning activities that are normally selected by the researcher for a typical science class. Indeed the literacy portion of the researcher's teaching has become more focussed and considered on a daily basis.

5.4 Recommendations for Future Research

The three intervention strategies employed in this research, although separately considered, were intertwined in their implementation. The F.R.E.S.H. scaffold and modelling of writing led into the discussions of peer/self-feedback. Indeed the five components of the F.R.E.S.H. scaffold were used to co-create the rubric that the students required to make their feedback more meaningful. The metacognitive instances in the journal were reflections of how useful the students found the scaffold or peer assessment to be in improving their writing. It is difficult to separate their relative worth from the overall improvement process. Certainly, together, they proved to be an effective intervention that developed students' abilities with scientific writing.

John Hattie's (1992) research into the effect sizes of various "educational innovations" that are designed to improve student outcomes indicates that, although most have merit, some innovations are better than others. This begs the question – of the three strategies employed in this research project, which had the largest effect? The researcher's opinion is that, in order of effect size, the largest effect was held by the introduction of the writing scaffold and modelling of its use. The second most effective strategy was considered to be the peer/self-assessment which leaves the metacognitive reflections as the least effective. It is probably not coincidental that this is also the order in which they were introduced to the students in the research project.

Future research along these lines could involve a comparative analysis of the three strategies with classes of similar ability levels. Each strategy could be employed in a separate class and quantitative measurements of writing skill taken across common assessment tasks. The rationale for such research could be in response to the findings for the fourth research question in this thesis concerning the problems that arose as a result of the implementation of the improvement strategies. The time allocated to teach the three strategies effectively might be too demanding for the curriculum in a particular course. The researcher's mind is drawn in particular to the time-sensitive science courses taught in the final years of secondary school. Once the most effective strategy is identified, it might be selected by itself for interventions in time-sensitive courses.

A longitudinal study of the Year 6 students who participated in this research might also provide some interesting results. Such a study might give insight into the amount of retention of the writing strategies in their subsequent years of Science education. Another measure of the effectiveness of this project could be to compare the achievement scores of the research participants in scientific writing with those students who didn't participate in the study. This, however, presents an ethical dilemma of equity for those students who did not participate in the study if there was a difference in achievement level in favour of the participating class. Interviews with selected students from the research group in later years might provide insight as to the longevity of the strategies. Some questions might include:

- Does the use of the F.R.E.S.H scaffold persist as a guide for students' writing of conclusions or does it become 'second-nature'?
- Do students continue to ask their peers for feedback on their writing, or assess its worth before submission (self-assessment)?
- Do students regularly consider the strategies they use for writing and have they found others (metacognition)?
- Have students used these strategies for other areas of their learning, even across disciplines?
- Have students shared the strategies with others?
- Do students feel that the strategies learnt in this research project have given them an advantage over non-participants?

It is thought that answers to these questions would develop an understanding of the three strategies employed in this study in terms of the long-term effects without causing undue resentment from non-participants in this study.

In hindsight, a few changes would have been made to this study. The interviews were conducted with only eight of the 19 students before and after the implementation of the intervention strategies. The interview 'selection criteria' should have been more flexible. The non-interviewed students' journals turned up interesting items that may have provided more insight into their self-efficacy such as Ralph's comment, "I think I have drop [sic], I just do not understand Science. It is Just does not like me. [sic]". An interview with Ralph might have clarified those comments. Also, the boy mentioned in the teacher journal who had a negative outburst in class was not followed up. Unfortunately those students' were not part of the interview group. It would have been good to clarify what their intent was and explore it a little more deeply. The researcher needed to be more reflexive in this part of the research and was too bound to the research instruments that had been put in place.

On the topic of reflexivity, the researcher's relative inexperience with conducting interviews meant that some topics were not explored (in retrospect) to their full extent. Some difficulty was experienced with the students' responses when the questions were too sophisticated or misunderstood. It would have been good to get more responses surrounding how they felt about writing in terms of self-efficacy. This area seemed hard to explore with the students. Often their responses would be vague and trail off or they would be a collection of thoughts that they thought would answer the question which Partington (2001, p. 34) identifies as a potential problem with interviews. They would soon become uncomfortable and it was hard to know how hard to press the idea without causing them undue stress. Often the subject would just be changed and the idea would remain unexplored. Being able to think about the research questions and catch small snippets of interest in the midst of the interviews was found incredibly difficult by the researcher. Follow-up interviews following the transcription of the initial interviews were considered but not performed to the researcher's regret.

Furthermore, it was noticed after the fact that the teacher journal as a data collection instrument had some flaws. Although for the most part the journal was written within one or two days of the lessons which involved content that was associated with the research, the events and interactions in the lesson were still based on the recollection of the researcher. These memories could be clouded by the emotional and mental state of the researcher at the time and potentially not be completely accurate representations of the conversations and events that transpired in the lessons. On reflection, a better record could have been kept by recording the lessons using video and leaving the teacher journal as a more personal and interpretive view of the lessons.

It was also clear that, whilst the majority of the class improved by leaps and bounds, this was not enjoyed by all. Some of the students only progressed by one level on the Conclusion Writing Continuum while others had a range of up to four levels. On further reflection, it is thought that presenting the F.R.E.S.H. scaffold as an external reference for the students may not have been enough for those student who were at a very low starting point. A "framed, or fill-in-the-blank, conclusion", as proposed by Rutherford (2007, p.n13) based on the work done by Terry (2001), for students with learning disabilities may have provided better results for those students. It is thought that a differentiated approach to the scaffolding may have been more effective, by engaging students in their Zone of Proximal Development (Vygostsky, 1978).

Over the course of the project, time pressures that were mentioned by the researcher in the teacher journal were becoming felt. As a result, there was only time for two opportunities for self-assessment. The time constraints imposed by the curriculum meant that the focus on conclusion writing had to be shifted by the end of allotted time for the research project. It is the researcher's opinion that more time may have been required to do justice to this portion of the feedback strategy. Future researchers might be mindful of this and the other points above.

5.5 Summary

In this chapter, the research questions were addressed in the first section. It was found that the intervention strategies had made the participating students believe that they had improved and were better scientific writers. Their self-efficacy towards writing had improved and there were many opportunities to realise success and exert control over their ability to write a scientific conclusion. Quantitative results showed a distinct improvement in the class' writing when measured against the Conclusion Writing Continuum. Several problems with the intervention strategies were discussed, including the initial trouble that students had with peer/self-assessment and the time taken to implement the strategies.

The second section described the implications of the research project. The students benefited the most from the research and are likely to carry the improved scientific writing skills with them into the future. The participating school gained higher-achieving students and insight into effective strategies for improving writing in science. The researcher, as a participant of the study, developed his own teaching skills and changed his views on the strategies employed. The research confirms the findings of many other studies (such as Rutherford, 2007, Hand, Wallace & Yang, 2004, LaConte & Berry, 2006, Brandt, 1971, Rice, 1998, Butler & Nesbit, 2008) and builds on professional knowledge in this field.

Finally the issues that were faced in the research and possibilities for future research were described in the final section of this chapter. A future comparison of the relative worth of the three intervention strategies was suggested, as well as a longitudinal study of the effects of this action-research in future years. Modifications to this research project were suggested in retrospect including changes to the interview process and increased scaffolding for weak writers.

This chapter concludes the thesis.

REFERENCES

Akerson, V., & Young, T. (2005). Science the "write" way. *Science and Children, 43* (3), 38–41

Anderson, G. (1998). *Fundamentals of educational research* (2nd ed.). New York: Routledge-Falmer.

Anderson, D., Nashon, S. M. & Thomas, G. P. (2009), Evolution of research methods for probing and understanding metacogntiion, *Research in Science Education, 39*, 181-185

Arends, R. (2001). *Learning to teach* (5th ed). New York: McGraw-Hill

Australian Curriculum Assessment and Reporting Authority. (2015). *MySchool* (Retrieved February 5, 2015 from the MySchool website: http://www.myschool.edu.au/)

Australian Curriculum Assessment and Reporting Authority. (2015). *Science Scope And Sequence.pdf* (Retrieved February 23/2015 from the Australian Curriculum Website: http://www.australiancurriculum.edu.au/australian%20curriculum.pdf?Type=0&s=S&e=ScopeAndSequence)

Australian Curriculum Assessment and Reporting Authority. (2016). *Science rationale*. (Retreived February 14/2016 from the Australian Curriculum Website: http://www.australiancurriculum.edu.au/Science/Rationale)

Baker, W. P., Barstack, R., Clark, D., Hull, E., Goodman, B., Kook, J., Kraft, K., Ramakrishna, P., Roberts, E., Shaw, J., Weaver, D., & Lang, M. (2008). Writing-to-learn in the inquiry-science classroom: Effective strategies from middle school science and writing Teachers, *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 81 (3), 105-108

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning, *Educational Psychologist*, *28* (2), 117-148

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behaviour* (pp71-81). New York: Academic Press

Bailey, D. M., & Jackson, J. M. (2003). Qualitative data analysis: Challenges and dilemmas related to theory and method. *The American Journal of Occupational Therapy,* 57 (1), 57-65

Bennett, B. (2011). *The problem with writing.* Retrieved February 23, 2013, from the Educator, Learner Website: www.brianbennett.org/blog/the-problem-with-writing/

Butler, M. B. & Nesbit, C. (2008). Using science notebooks to improve writing skills and conceptual understanding, *Science Activities: Classroom Projects and Curriculum Ideas, 44* (4), 137-146

Brandt, W. W. (1971). Practice in critical reading as a method to improve scientific writing. *Science Education, 55* (4). 451-455

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

Clarke, M. C., & Sharf, B. F. (2007). The dark side of the truth(s). Ethical dilemmas in researching the personal. *Qualitative Inquiry, 13* (3), 399-416

Connelly, F. M., & Clandinin, D. J. (1990). Stories of experience and narrative inquiry. *Educational Researcher, 19* (5), 2-14

Conner, L. N. (2007). Cueing metacognition to improve researching and essay writing in a final year high school biology class. *Science Education,* 37 (1), 1-16

Cresswell, J. W. (2005) *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (2nd ed.). Upper Saddle River, NJ: Pearson

Dweck, C. (2000). *Self-theories: Their role in motivation, personality, and development. Essays in social psychology*. Florence: Psychology Press, Taylor and Francis Group

Einarsdottir, J. (2007). Research with children: methodological and ethical challenges. *European Early Childhood Education Research Journal, 15* (2), 197-211

Erickson, F. (1986). Qualitative methods in research on teaching. In M. C. Wittrock (Ed.), *Handbook of Research on Teaching, 3rd Edition*, New York, Macmillan.

Etherington, K. (2007). Ethical research in reflexive relationships. *Qualitative Inquiry*, *13* (5), 599-616

Flavell, J. H. (1979). Metacognition and Cognitive Monitoring: A New Area of Cognitive—Developmental Inquiry. *American Psychologist, 34* (10), 906-911

Gallagher, J. J., Parker, J., & Ngwenya, J. (1999). *Embedded assessment and reform in science teaching and learning*. East Lansing, MI: Michigan State University.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.

Glogger, I., Schwonke, R., Holzapfel, L., Nuckles, M., & Renkl, A. (2012). Learning strategies assessed by journal writing: Prediction of learning outcomes by quantity, quality and combinations of learning strategies. *Journal of Educational Psychology*, *104* (2), 452-468

Hand, B., Wallace, C.W., & Yang, E. (2004). Using a science writing heuristic to enhance learning outcomes from laboratory activities in seventh-grade science: quantitative and qualitative aspects, *International Journal of Science Education*, *26* (2), 131-149

Hattie, J. 1992. Measuring the effects of schooling. *Australian Journal of Education 36* (1), 5-13

Hohenshell, L., Woller, M., & Wallace, S. (2013). On the road to science literacy: building confidence and competency in technical language through choral repetition. *Journal of College Science Teaching*, *42* (6), 38-43

Isaac, S., & Michael, W. B. (1987). *Handbook in research and evaluation* (2nd ed.). San Diego, CA: EdITS Publishers.

Kayaoglu, M. N. (2015). Teacher researchers in action research in a heavily centralized education system. *Educational Action Research*, *23* (2), 140–161

Keys, C. W., Hand, B., Prain, V., & Collins, S. (1999) Using the science writing heuristic as a tool for learning from laboratory investigations in secondary science, *Journal of Research in Science Teaching*, *36* (10), 1065–1084

Klassen, S. (2006). Contextual assessment in science education: Background, issues, and policy. *Science Education, 90* (5), 820-851

Kocakülah, M. S. (2010). Development and application of a rubric for evaluating students' performance on Newton's laws of motion. *Journal of Science Education and Technology*, *19* (2), 146-164

Krajcik, J. S., & SutherLand, L. M. (2010). Supporting students in developing literacy in science. *Science*, *328*, 456 -459

Kyle, W. C. (Ed.) (1997). Editorial: Action Research. *Journal of Research in Science Teaching, 34* (7), 669-671

LaConte, J., & Berry, J. (2006). Writing in science: A collaborative approach to improve practice. *Science Scope, 30* (1), 63-65

Lawrenz, F. Huffman, D. & Welch, W. (2000). The science achievement of various subgroups on alternative assessment formats. *Science Education*, *85* (3), 279-290

Lee, O., & Luykx, A. (2005). Dilemmas in scaling up innovations in elementary science instruction with nonmainstream students. *American Educational Research Journal, 42* (3), 411–438

Lloyd, D. (2007). Exploring students' future images. In P. C. Taylor & J. Wallace (Eds.), *Contemporary qualitative research: Exemplars for science and mathematics educators,* (pp59-68). Dordrecht, The Netherlands: Springer.

Mason, J. (1994) Linking qualitative and quantitative data analysis. In A. Bryman & R. G. Burgess (Ed.), *Analyzing qualitative data* (pp89-110). Routledge: London.

Microsoft. (2013). *Microsoft Excel* [computer software]. Redmond, Washington: Microsoft.

Mulholland, J., & Wallace, J. (2003). Crossing borders: Learning and teaching primary science in the pre-service to in-service transition. *International Journal of Science Education, 25* (7), 879-898

Norris, S. P., & Phillips L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education*, *87* (2), 224-240

Partington, G. (2001). Qualitative research interviews: Identifying problems in technique. *Issues in Educational Research, 11* (2), 32-44

Peha, S. (2003). *Writing across the curriculum.* Retrieved February 23, 2013, from Teaching That Makes Sense Web site: www.ttms.org/PDFs/06%20Writing%20Across%20the%20Curriculum%20v001%20(Full).pdf

Polkinghorne, D. E. (1995). Narrative configuration in qualitative analysis. *Qualitative Studies in Education*, *8* (1), 5–22

Polkinghorne, D. E. (2007). Validity issues in narrative research. *Qualitative Inquiry*, *13* (4), 471-486

Porter, R., Guarienti, K., Brydon, B., Robb, J., Royston, A., Painter, H., Sutherland, A., Passmore, C., & Smith, M. H. (2010). Writing better lab reports. *The Science Teacher, 77* (1), 43-48

Rivard, L.P. (1994). A review of writing to learn science: Implications for practice and research. *Journal of Research in Science Teaching, 31*, 969-983

Rice, R. E. (1998). Scientific writing - A course to improve the writing of science students. *Journal of College Science Teaching, 27* (4), 267-272

Ruiz-Primo, M. A., & Shavelson, R. J. (1996). Rhetoric and reality in science performance assessment: An update. *Journal of Research in Science Teaching*, *33*, 1045–1063

Russell, B. H. (1988). *Research methods in cultural anthropology*, Newbury Park, CA: Sage.

Rutherford, S. (2007) Using a laboratory conclusion rubric. *Science Activities: Classroom Projects and Curriculum Ideas, 43* (4), 9-14

Simon, S. (1992). Curriculum and assessment innovation in science. *Research in Science Education, 22* (1), 358-366

Shaw, J. M. (1997). Threats to the validity of science performance assessments for English language learners. *Journal of Research in Science Teaching, 34*, 721–743

Shepardson, D. & Britsch, S. (2001). Tools for assessing and teaching science in elementary and middle school. In D. Shepardson (Ed.), *Assessment in science: A guide for professional development and classroom practice* (pp119-147). Dordrecht, Netherlands: Kluwer Academic Publishers.

Snowman, J., & Biehler, R. (2000). *Psychology applied to teaching*. Boston: Houghton Mifflin

Statistics How To. (2015). *Pearson correlation: Definition and easy steps for use.* (Retrieved 12 March 2015 from the Statistics How To website:http://www.statisticshowto.com/what-is-the-pearson-correlation-coefficient/)

Terry, B. M. (2001). *Turn students with learning disabilities into writers*. (Retrieved October 19, 2006 from Parent Pals Website: http://www.parentpals.com/gossamer/pages/Detailed/889.html)

Treagust, D. F., Jacobowitz, R, Gallagher, J. L., & Parker, J. (2001) Using assessment as a guide in teaching for understanding: A case study of a middle school science class learning about sound. *Science Education, 85* (2), 137-157

Vygostsky, L. S. 1978. *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press

Warwick, P., & Maloch, B. (2003). Scaffolding speech and writing in the primary classroom: A consideration of work with literature and science pupil groups in the USA and UK. *Reading*, *37* (2), 54-63

Wiggins, G. (1998). *Educative assessment*. San Francisco: Jossey-Bass.

Willis, J. W. (2007). *Foundations of qualitative research: Interpretive and critical approaches*. Thousand Oaks, CA: Sage

Wollman-Bonilla, J. E. (2000). Teaching science writing to first graders: Genre learning and recontextualisation. *Research in the Teaching of English, 35* (1), 35–65

APPENDICES

Appendix A: Information Letter for Participants

Curtin University

Dear Parent/Guardian,

I am your child's Science Teacher, Duncan Wood. I am currently doing a Master of Philosophy in Science Education with Curtin University of Technology. I am interested in how students write about scientific concepts and their scientific reasoning about the experiments they perform in class.

The title of the project is: *Improvement of students' scientific writing in a middle-years Science classroom*.

<u>Aims and description of the project</u>

Writing in Science is something that is often challenging for students. Often the requirements for a successful piece of writing in science can be different than those experienced in other classes such as English. Difficulty can stem from a variety of sources such as the use of technical language, the level of scientific reasoning and understanding of scientific principles. This semester your son/daughter will be developing different strategies to improve their writing. These strategies are based on research conducted by science educators from around the globe. They are as follows:

1) The use of a scaffolded framework for writing assessments that will be developed with the students in the form of a series of 'success criteria'. The framework will be visible during class time and during assessments in the form of a large poster in plain sight. The framework will be referred to during practice tasks and for modelling examples with the students.
2) The second strategy will be the use of peer/self-assessment with respect to the framework developed with the students in strategy one. A small amount of class-time will be spent developing guidelines and a rubric for students to follow.
3) The third strategy employed will be the use of a journal for the students to reflect on their experiences with writing assessments, hopefully identifying what was productive and which strategies help.

Requirements of your son/daughter/ward

With your permission I would like to conduct two interviews with your child/ward about their experience with science writing and which strategies they found most helpful. The interviews will take around 5-10 minutes each and will be at the start and end of term 2 taking place at a convenient time outside of classes. I will record their responses and I will also be taking copies of their written work that they perform in class. Interview recordings and transcripts will be available for you and your child to review at any stage of the study. You or your child/ward may withdraw support for the research at any time without giving a reason, in which case the interview recordings and transcripts will not be used.

Confidentiality and security of information

Only my supervisor, Doctor Rekha Koul of Curtin University and I, will have access to the information. Results of this research will be published as a thesis. However, it will be made completely anonymous and there will be no way of identifying the students in any published material. Participation is completely voluntary; you or your child are at liberty to withdraw at any time without prejudice or negative consequences; If you choose to decline, your child will in no way be affected in or out of class. However, participating students will benefit from reflecting on their work in a more formal way for the loss of a small amount of time outside of class.

Should you wish to contact me to discuss any of the details of the research I am contactable at the school on Telephone: [Research school phone number] or by Email [Researchers email address]

Alternatively you can contact my supervisor Doctor Rekha Koul on Telephone:(08) 9266 4074 or by Email: R.Koul@curtin.edu.au

Also, should participants wish to make a complaint on ethical grounds the contact details of the Human Research Ethics Committee (Secretary) are as below

Phone: +61 (08) 9266 2784 or by Email: hrec@curtin.edu.au

or in writing C/- Office of Research and Development, Curtin University of Technology, GPO Box U1987, Perth WA 6845);

The project has been approved by the Curtin University Human Research Ethics Committee, the approval number is SMEC-39-12

Yours sincerely,

Duncan Wood

Please fill in your details and return to Mr Wood

Title of project: Improvement of students' scientific writing in a middle-years Science classroom.

Researcher: Mr Duncan Wood

I have been informed of and understand the purpose of the study. I have been given an opportunity to ask questions. I understand that I or my child/ward can withdraw at any time without prejudice. Any information which might potentially identify me or my child/ward will not be used in published material. I agree to allow my child/ward to participate in the study as outlined in the information sheet.

Name_________________________________

Name of child/ward _______________________________

Signed ______________________________ Dated _____/______/__________

Appendix B: Strong and Weak Work Sample Tasks

Acid and Limestone

Larry hypothesises that a reaction between an acid and a piece of limestone is going to create some hydrogen gas. He has seen in Year 6 Chemistry that acid labels tend to have the letter H for Hydrogen in them and that acids often create Hydrogen gas when you add them to metals. He knows that a pop test is a good way to determine if Hydrogen gas is being formed.

Larry takes several chunks of limestone (also known as Calcium Carbonate $CaCO_3$) from different locations in a limestone cave and brings them to a lab. He sets up 5 test tubes with the chunks in the bottom of them and adds 10mL of Hydrochloric acid (HCl) to each of them. He then tapes another test tube over the top of them to collect the gas. He notices bubbles forming rapidly on the surface of the limestone chunks. Lighting a match, he takes each gas filled tube off after 1 minute of collecting and places it over the match as in the pop test. Each time he does it the match goes out straight away and there is no pop. He decides to record exactly how long it take the match to extinguish and gets 0.5sec, 0.6sec, 0.5sec, 0.4sec and 0.5sec for the five test tubes. He also noticed that the limestone seemed to be eaten away.

<u>Conclusion 1</u>

Larry's hypothesis was incorrect, acid and limestone do not create hydrogen gas. A gas was seen because of the bubbling (in this sense he was correct) but it wasn't hydrogen. If hydrogen gas had been created there would have been a pop when he held the lit match under the test tube, instead it just went out. The gas that was formed might have been something different. The results seem fairly reliable as it happened each time he did the experiment and consistently took about 0.5 seconds on average to extinguish the flame. Larry could try using different type of acid to see if hydrogen gets produced with them.

Conclusion 2

Larry should have done more tests, maybe he would have got the gas he was looking for. Sometimes chemical tests can be faulty so he shouldn't lose heart. There might have been a leak in the tape so the gas escaped. Overall he seemed to control things that could affect the experiment like the amount of acid.

Conclusion 3

The hypothesis was incorrect. Hydrogen only forms when an acid reacts with metal. More than likely it was Carbon Dioxide created during the experiment because it is well known to extinguish flames. As Carbon dioxide is heavier than air it probably floated down out of the test tube onto the flame and put it out. Also, Carbon and Oxygen are present in the formula for Limestone/Calcium Carbonate (CaCO3), so the acid has probably forced the CO2 out of the limestone. Further testing should include tests for other gases other than hydrogen (like CO_2). The amount of limestone needs to be controlled more precisely (chunk is not a precise measurement).

The Rusty Nail Experiment

An experiment is performed to determine how much rust is created on nails in different environments. Nails are put into 3 test tubes, one filled with water, one half-filled with water and one with no water. Stoppers are put in and they are left for a week. A hypothesis is made that the one with the most water will have the most rust. The test is performed by four different groups using the same method.

A week later, the rust is scraped from each nail and measured (all masses in grams)

Test tube	Group 1	Group 2	Group 3	Group 4
Air Only	0.01	0.01	0.01	0.01
Air/Water Combo	0.05	0.04	0.05	0.04
Water only	0.02	0.03	0.03	0.02

Select the best sentences from the list below to create a high level hypothesis

Hypothesis sentence

The hypothesis was incorrect, the most rust was created in a combination of air and water

The water one didn't have as much rust, so the hypothesis was incorrect

When nails are put in water they don't rust very much.

Evidence sentence

All the nails had some rust, some more than others.

The order of nail environments from most rusty to least rusty was air/water, water only, air only.

In one test, the air/water combo had 0.05g of rust, the water one had 0.02g and the air had only 0.01g of rust.

Theory sentence

Rust happens when you leave something in the rain for a while, so air and water are needed

Iron rust occurs when oxygen in the air bonds to the iron. Water just allows this to happen easier which is why the air/water combo worked the best.

Nails get rusty all the time, but water and air help it along. Corrosion is another name for rust.

Reliability sentence

The nails weren't all exactly the same size so measurement of the amount of rust may have depended on this. The nails weren't all that different so it probably doesn't change the conclusion. All the groups got the same order which means the experiment was reliable.

The experiment was fine and the results were reasonable

One group's data was pretty much the same as another group's data so it was a reliable experiment

Further Experimentation sentence

They should have put the nails into wood to see if it makes a difference

As the reaction occurs between the iron and air in the presence of water, different coatings for the nail should be tested to see how waterproof they can make them. The size of the nails should be more tightly controlled in the next test.

We should test how much rust is created on big nails versus little nails in the same environments, air/water combo, air only, and water only. Then we could compare the results with this experiment.

Now copy the best sentences into your book.

Appendix C: Conclusion Writing Tasks

Jellyfish Task

Write a conclusion to the experiment below.

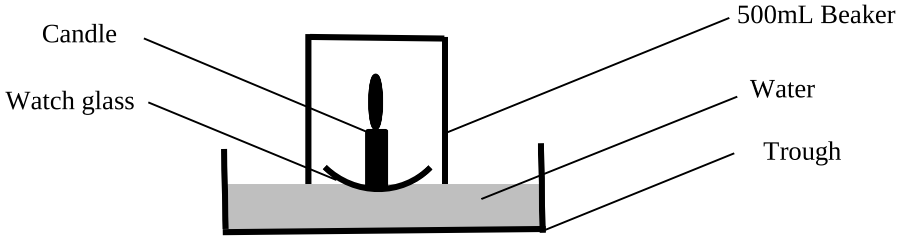
A chemist wants to know if adding vinegar to a jellyfish sting is beneficial. She hypothesises that jellyfish stings are alkaline and that vinegar should neutralise it.

Pain reported on a scale of 1-10 of a jellyfish sting

Before applying vinegar	After applying vinegar
5	3
7	2
9	2
10	6
8	8

Candle Task

Set up the following apparatus



What is going to happen when the candle is lit, and the beaker placed over it? Write a hypothesis.

Conduct the experiment twice and then write a conclusion based on your observations.

Water Quality Task

You will be given 6 samples of water taken from around [home town of the research school] and be asked to test each sample for water quality.

- Before starting, consider the location that each sample was taken from and hypothesise about the quality of the water sample.
- Next, perform chemical testing of water quality using nitrate/phosphate testing strips and pH and temperature probes
- Graph your results on Excel
- Write a conclusion for the experiment

Hydrogen Task

A Year 6 student is interested in how quickly hydrogen gas is given off by a reaction between acid and magnesium. She performs an experiment where different concentrations (strength) of acid are added to 5 grams of magnesium powder in a conical flask. She times it until fizzing stops. She **hypothesises** that the stronger the acid, the more quickly the hydrogen will be produced. Is she right?

Concentration of acid (M)	Trial 1 (s)	Trial 2 (s)	Trial 3 (s)
0	49	63	57
1	28	30	31
2	20	22	21
3	16	17	16
4	10	9	25
5	7	6	6
6	5	6	4

Write a conclusion in the space below for the experiment.