1991

Monitoring Standards Of Science Investigation Skill Attainment By Tongan Secondary Science Students

Fisi'ihoi Mone

Edith Cowan University

Recommended Citation


This Thesis is posted at Research Online.
https://ro.ecu.edu.au/theses/1457
You may print or download ONE copy of this document for the purpose of your own research or study.

The University does not authorize you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following:

- Copyright owners are entitled to take legal action against persons who infringe their copyright.

- A reproduction of material that is protected by copyright may be a copyright infringement. Where the reproduction of such material is done without attribution of authorship, with false attribution of authorship or the authorship is treated in a derogatory manner, this may be a breach of the author’s moral rights contained in Part IX of the Copyright Act 1968 (Cth).

- Courts have the power to impose a wide range of civil and criminal sanctions for infringement of copyright, infringement of moral rights and other offences under the Copyright Act 1968 (Cth). Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.
USE OF THESIS

The Use of Thesis statement is not included in this version of the thesis.
Monitoring Standards of Science Investigation Skill Attainment by Tongan Secondary Science Students

By

Fisi'ihoi Mone B.Sc., Post-Grad.

A thesis submitted in partial fulfilment of the requirement for the award of

Master of Education

at the Faculty of Education
Edith Cowan University
June 1991
ABSTRACT

The main purpose of this study was to evaluate the science investigation skill attainment of Tongan Form 5 (16 years of age) General Science students. Benchmark statements were developed to describe the range of science investigation skills and standard of performance that should be expected of Tongan Form 5 General Science students. A written test of science process skills and a practical test of science apparatus skills, were developed to assess the level of attainment of science investigation skills by students who have completed Form 4 and Form 5 General Science in Tonga. The instruments were piloted twice in Western Australian schools, revised, piloted in Tonga and then administered to students at ten high schools in Tonga.

From the written test of science process skills, it was found that more than 60% of the Form 5 students had not attained the benchmark standards. The students performed best on questions regarding collecting and communicating information, and worst on the questions relating to problem analysis, planning and control of variables. From the practical test of science apparatus skills, it was found that more than half of the Form 5 students had not attained the benchmark standards related to using laboratory equipment like a thermometer, Bunsen burner, triple-beam balance, and measuring cylinder.
DECLARATION

I certify that this thesis does not incorporate, without acknowledgement, any material previously submitted for a degree or diploma in any institution of higher education and that, to the best of my knowledge and belief, it does not contain any material previously published or written by another person except where due reference is made in the text.

Fisi'ihoi Mone
ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to those who have made the completion of this project possible.

Special thanks are extended to the project's supervisor, Dr Mark Hackling, from whom guidance, assistance, encouragement and constructive criticism were always given.

Sincere thanks are also extended to Dr Pat Garnett, the project's second supervisor, for his assistance, constructive criticism and time.

I would also like to acknowledge the assistance provided by Dr Barry Sheridan in the evaluation of the psychometric properties of the test instruments, and Paul Carreno in the analysis of the results of the tests.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>i</td>
</tr>
<tr>
<td>DECLARATION</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td><strong>CHAPTER 1</strong> INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Context</td>
<td>1</td>
</tr>
<tr>
<td>The Research Problem</td>
<td>2</td>
</tr>
<tr>
<td>Significance of the Study</td>
<td>3</td>
</tr>
<tr>
<td>Purpose and Research Questions</td>
<td>4</td>
</tr>
<tr>
<td>Definition of Key Terms</td>
<td>5</td>
</tr>
<tr>
<td><strong>CHAPTER 2</strong> LITERATURE REVIEW</td>
<td>6</td>
</tr>
<tr>
<td>Introduction</td>
<td>6</td>
</tr>
<tr>
<td>The Role of Laboratory Work</td>
<td>6</td>
</tr>
<tr>
<td>Acquisition of Skills</td>
<td>9</td>
</tr>
<tr>
<td>Assessment of Skills</td>
<td>15</td>
</tr>
<tr>
<td><strong>CHAPTER 3</strong> METHODOLOGY</td>
<td>23</td>
</tr>
<tr>
<td>Introduction</td>
<td>23</td>
</tr>
<tr>
<td>Instrument Development</td>
<td>23</td>
</tr>
<tr>
<td>Subjects</td>
<td>31</td>
</tr>
<tr>
<td>Procedure</td>
<td>34</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>36</td>
</tr>
<tr>
<td><strong>CHAPTER 4</strong> EVALUATION OF THE TEST INSTRUMENTS</td>
<td>39</td>
</tr>
<tr>
<td>Introduction</td>
<td>39</td>
</tr>
<tr>
<td>Written Test of Science Process Skills</td>
<td>39</td>
</tr>
<tr>
<td>Practical Test of Science Apparatus Skills</td>
<td>49</td>
</tr>
<tr>
<td><strong>CHAPTER 5</strong> RESULTS</td>
<td>51</td>
</tr>
<tr>
<td>Introduction</td>
<td>51</td>
</tr>
<tr>
<td>Performance of Students on the Written Test of Science Process Skills</td>
<td>51</td>
</tr>
<tr>
<td>Performance of Students on the Practical Test of Science Apparatus Skills</td>
<td>55</td>
</tr>
<tr>
<td>Gender Differences in Performance of Form 5 Students on the Tests</td>
<td>60</td>
</tr>
<tr>
<td>Performance of Form 5 Students in Relation to Benchmark Standards</td>
<td>64</td>
</tr>
<tr>
<td>Summary of the Main Findings</td>
<td>67</td>
</tr>
<tr>
<td>Chapter</td>
<td>Section</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Chapter 6</td>
<td>DISCUSSION</td>
</tr>
<tr>
<td></td>
<td>Introduction</td>
</tr>
<tr>
<td></td>
<td>Performance of Students on the Written Test of Science Process Skills</td>
</tr>
<tr>
<td></td>
<td>Performance of Students on the Practical Test of Science Apparatus Skills</td>
</tr>
<tr>
<td></td>
<td>Gender Differences in Performance of Form 5 Students on the Tests</td>
</tr>
<tr>
<td></td>
<td>Assessment Procedures in Tonga</td>
</tr>
<tr>
<td>Chapter 7</td>
<td>CONCLUSIONS AND IMPLICATIONS</td>
</tr>
<tr>
<td></td>
<td>Introduction</td>
</tr>
<tr>
<td></td>
<td>Conclusions of the Study</td>
</tr>
<tr>
<td></td>
<td>Implications for Teaching</td>
</tr>
<tr>
<td></td>
<td>Implications for Further Research</td>
</tr>
<tr>
<td>References</td>
<td></td>
</tr>
<tr>
<td>Appendix 1</td>
<td>Benchmark Statements</td>
</tr>
<tr>
<td>Appendix 2</td>
<td>Objectives Tested in the Written Test of Science Process Skills</td>
</tr>
<tr>
<td>Appendix 3</td>
<td>Test of Science Process Skills</td>
</tr>
<tr>
<td>Appendix 4</td>
<td>Marking Key for the Written Test of Science Process Skills</td>
</tr>
<tr>
<td>Appendix 5</td>
<td>Objectives Tested in the Practical Test of Science Apparatus Skills</td>
</tr>
<tr>
<td>Appendix 6</td>
<td>Test of Science Apparatus Skills</td>
</tr>
<tr>
<td>Appendix 7</td>
<td>Marking Key for the Practical Test of Science Apparatus Skills</td>
</tr>
<tr>
<td>Appendix 8</td>
<td>Equipment and Materials List for the Practical Test of Science Apparatus Skills</td>
</tr>
<tr>
<td>Appendix 9</td>
<td>Administration Guide for the Practical Test of Science Apparatus Skills</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Number of Students and Schools Selected for the Study ........................................... 32</td>
</tr>
<tr>
<td>2</td>
<td>Fit Order of the Multiple Choice Items in the Second Trial of the Written Test ........ 40</td>
</tr>
<tr>
<td>3</td>
<td>Fit Order of the Total Items in the Second Trial of the Written Test ...................... 43</td>
</tr>
<tr>
<td>4</td>
<td>Fit Order of the Multiple Choice Items in the Written Test of Science Process Skills 46</td>
</tr>
<tr>
<td>5</td>
<td>Fit Order of All Items in the Written Test of Science Process Skills ..................... 47</td>
</tr>
<tr>
<td>6</td>
<td>Fit Order of the Total Items in the Practical Test of Science Apparatus Skills ........ 49</td>
</tr>
<tr>
<td>7</td>
<td>Maximum Possible Scores and Mean Scores for Form 3 and Form 5 students on Each Item of the Written Test .......................................................... 52</td>
</tr>
<tr>
<td>8</td>
<td>Maximum Possible Score, Mean and Standard Deviation for Form 3 and Form 5 Students on the Skill Areas of the Written Test ........................................... 54</td>
</tr>
<tr>
<td>9</td>
<td>Maximum Possible Score, Mean Total Scores and Standard Deviations for Form 3 and Form 5 students on the Written Test ................................................. 56</td>
</tr>
<tr>
<td>10</td>
<td>Maximum Possible Score and Mean Score for Form 3 and Form 5 students on Individual Items of the Practical Test ................................................................. 56</td>
</tr>
<tr>
<td>11</td>
<td>Maximum Possible Score, Mean and Standard Deviation for Form 3 and Form 5 students on the Skill Areas of the Practical Test ............................................. 58</td>
</tr>
<tr>
<td>12</td>
<td>Maximum Possible Score, Mean Total Scores and Standard Deviations for Form 3 and Form 5 students on the Practical Test ...................................................... 59</td>
</tr>
<tr>
<td>13</td>
<td>Mean Scores for Form 5 Boys and Girls on Each Item of the Written Test .................... 60</td>
</tr>
</tbody>
</table>
14 Mean and Standard Deviation for Form 5 Boys and Girls on Each Skill Area of the Written Test 62
15 Mean and Standard Deviation for Form 5 Male and Female Students on Each Skill Area of the Practical Test 63
16 Percentage of Form 5 Students Below, At, and Above Benchmark Standards in the Skill Areas of the Written Test 65
17 Percentage of Form 5 Students Below, At, and Above Benchmark Standards in the Skill Areas of the Practical Test 66
CHAPTER 1

Introduction

Context

The current General Science curriculum taught in Tongan secondary schools was developed by the Tongan Ministry of Education to replace the New Zealand General Science program. It covers all levels from Form 1 (12 years of age) up to Form 5 (16 years of age). It was adapted from the New Zealand General Science course and is still at an early stage of its development.

Practical work is a compulsory part of this curriculum. However, science investigation skills, that is, the process skills and apparatus skills used in conducting science investigations, are only assessed through students' written reports of laboratory activities. There has been no literature written nor a study reported regarding the level of science investigation skill attainment of the Tongan secondary General Science students.
CHAPTER 1

The Research Problem

Since the termination of the New Zealand School Certificate and University Entrance examinations in the mid 1980s, the Tongan Ministry of Education has been faced with the task of developing the curriculum and assessment methods to take the place of the New Zealand ones. New curricula have been developed by making minor modifications to the New Zealand materials. The main objective has been to produce a program for Form 4 and Form 5 that will not only be relevant to Tonga, but also be recognized by overseas tertiary institutions. Students sit the Tongan School Certificate Examination based on Form 4 and Form 5 programs.

Although two of the seven major objectives in the Form 5 General Science program suggest application of scientific method and development of skills appropriate to science, the content of the curriculum package reveals that laboratory work is employed mainly as a vehicle for developing the knowledge component of the course. Many writers in science education now believe that the main benefit of practical work is the acquisition of skills that are related to the processes of science (Kirschner, 1989; Woolnough & Allsop, 1985). There is a need to assess the extent to which students in Tonga have acquired science investigation skills, and to make recommendations for revision of the curriculum materials.
Significance of the Study

This project can contribute to the development of the Tongan General Science curriculum in many respects. First, it will stimulate an increased awareness of the importance of teaching science investigation skills. Second, it will establish benchmark standards for science investigation skills on which future development of the course can be based. These will be accompanied by assessment instruments which can be used in the future to examine and monitor the developments of the course with regard to science investigation skills. Third, it will not only identify the laboratory skills included in the intended curriculum, but also those that students have attained after completing the Form 4 and Form 5 courses. This will help identify the areas of the program that need improvement.

This study will be the first of its kind to be conducted in Tonga. The results will not only be important to the development of the General Science course in Tonga, but will also form a base for further research in this area.
CHAPTER 1

Purpose and Research Questions

The purposes of this study are to establish benchmark standards for science investigation skills, to develop assessment instruments to evaluate students' attainment of these skills, and to make recommendations for changes to the science curriculum of Tonga.

More specifically this study will address the following research questions:

1. Which science investigation skills and standards of performance should Form 5 Tongan General Science students attain?

2. Which skills have students attained before and after completing the Tongan Form 4 and Form 5 General Science programs?

3. How effective are the Form 4 and Form 5 curricula in developing science investigation skills?

The subsidiary question that will also be addressed by this study is:

4. Do Form 5 male and female students differ in levels of skill attainment?
Definition of Key Terms

1. **Science investigation skills.** These are the skills which are involved in carrying out science investigations. They can be divided into two main types:

   (a) **Science process skills.** The intellectual skills which students use as they plan investigations, collect and interpret data, and formulate responses to questions.

   (b) **Apparatus skills.** The skills involved in using science apparatus. These involve the procedural knowledge, cognitive skills and manipulative skills associated with the use of particular apparatus.

2. **Tongan General Science Program.** A general science program prepared by the Tongan Ministry of Education for students in their seventh to eleventh years of formal education. It is a modified New Zealand program.
CHAPTER 2

Literature Review

Introduction

This chapter reviews the literature relevant to this study. The main purpose of this review is to establish the role of laboratory work in high school science, describe the processes of skill learning, and identify appropriate methods for assessing science investigation skills.

The Role of Laboratory Work

Studies conducted in Wales and England revealed that the most important of science teachers' goals for practical work for 11-13 year olds were those relating to the acquisition of practical skills (Beatty & Woolnough, 1982). Kirschner and Meester (in Kirschner, 1989) suggested three motives for practical work. The first of these motives is the development of process skills and laboratory techniques. The second is to develop an academic approach to working, that is, being a problem solving scientist, and the third is to 'experience' or 'get a feel for' phenomena.
Kirschner and Meester distinguished three major types of practical work: simulations, experimental seminars, and wet laboratories. Simulations are organized experiences in which reality is imitated. These allow the learner to experience reality in a less expensive and quicker way than wet laboratories. In experimental seminars students collectively perform an experiment or watch an experiment performed by an expert, followed by group discussions. Wet laboratories, which are the most common type of activity prescribed in science curricula, enable students to gain hands-on experience in a laboratory setting. Kirschner and Meester believed that simulations serve the development of process skills; experimental seminars serve the academic approach to working; and wet laboratories allow the students to experience phenomena and develop apparatus skills.

Kirschner and Meester also contended that using practical activities to illustrate or affirm theories, to enable discovery learning in the laboratory, and to allow students to distil insight and understanding from empirical work with phenomena, should not be the main focus of laboratory work. They advocated that these can only be achieved after the learner has acquired a broad critical knowledge of the particular subject matter through formal learning processes. Similarly, Woolnough and Allsop (1985) argued that the "imposition of theory on practical work has
a detrimental effect on the development of scientific investigation skills" (p. 38). Further, they pointed out that the imposition of practical work on theory also hinders the development of cognitive understanding. They stressed that practical work is best directed towards achieving goals that relate to the processes of science.

Science investigation skills are the skills that students use when they are involved in science practical work. They can be divided into two categories, science process skills and science apparatus skills. Science apparatus skills are the skills involved in using scientific apparatus. These skills involve procedural knowledge, cognitive skills and manipulative skills associated with the use of particular apparatus. Science process skills are "intellectual skills which students use in the classroom when they collect and interpret data" (Tobin & Capie, 1980, p. 590). Millar and Driver (1987) defined them as a "toolkit of strategies for tackling the task of gaining knowledge of the physical world" (p. 41). Lunetta, Hofstein and Giddings (1981) suggested that students' behaviour in conducting science investigations can be divided into four phases of activity. They are planning and design, performance, analysis and interpretation, and application. Within these phases many skills are applied to an investigation.
In the planning and design phase, the students are involved in formulating questions, predicting results, formulating hypotheses to be tested and designing experimental procedures. In the performance phase, the students are involved in conducting the experiments, manipulating materials, making decisions about investigative techniques and observing and recording data. In the analysis and interpretation phase, the students are involved in processing data, explaining relationships, developing generalizations, examining the accuracy of data, outlining assumptions and limitations, and formulating new questions based on the investigation. In the application phase students make predictions about new situations based on the results, and apply findings and techniques to new problems and variables (Lunetta et al., 1981).

**Acquisition of Skills**

Oslon (1973) postulated that the information gained through any activity was composed of two components namely knowledge and skill. This distinction, he said, between knowledge and skill corresponds to that suggested by Ryle (1949) between propositional knowledge (knowing that) and procedural knowledge (knowing how).
Oslon distinguished three modes of learning; direct experience, modelling or observational, and symbolically coded. He observed that the most common mode of learning used in classrooms had been the symbolically coded form, that is information transmitted through the media of speech, print, pictures and films. There were two reasons for this. First, the symbolic form had been perceived to be the most suitable one for children and second, teachers had been usually confronted with large groups of children. This mode of learning may serve well the acquisition of knowledge, but to acquire skills, the students require the direct experience mode of learning. This is one of the reasons why schools have been more successful in dealing with knowledge than developing skills.

Fitts (1964) proposed three phases in the acquisition of complex skills. First is the cognitive phase. The learner must acquire a certain knowledge of the substantive structure of a domain prior to the learning of a skill. The second is the intermediate or associative phase. At this stage, a dominant role is played by practice and feedback which helps eliminate the inappropriate characteristics of the new patterns of skill. This process goes on until the learner is capable of carrying out a complex skill as an uninterrupted unit instead of a series of simpler skills. The third phase, called the autonomous stage, is where the learner gains speed, control and coordination of the different subskills which make up the skill.
Anderson (1982) further developed the work of Fitts (1964) on learning cognitive skills. He proposed a framework for skill acquisition that includes two major stages. First is a declarative stage, which is characterised by the interpretation of facts about the skill domain during the performance of that skill. Second is a procedural stage which is characterised by the direct embodiment of the domain knowledge in procedures for performing the skill. In a computer based model of skill development called ACT (Adaptive Control of Thought), declarative knowledge is represented as a propositional network, while procedural knowledge is represented as productions (Anderson, 1976). The process by which the skill is transformed from the declarative stage to the procedural stage is called knowledge compilation. It has two subprocesses. The first is proceduralization, in which factual knowledge is embedded into productions. The second subprocess is composition. In this process, sequences of productions are collapsed into a single production. After the skill has been proceduralised, its application gains more speed. From this stage, further learning mechanisms operate on the skill, causing the productions to be more selective in their range of applications. These mechanisms include generalization, discrimination, and strengthening of productions.
Anderson (1983) suggested that "It is not possible to simply add a production in the way it is possible to simply encode a cognitive unit. Rather, procedural learning occurs only in executing a skill; one learns by doing" (p. 215).

Woolnough and Allsop (1985) advocated the importance of 'getting a feel for' or 'an awareness of' scientific phenomena. They called this 'tacit knowledge'. They argued that "When making a device, or solving a problem, they (students) will know what material to use and which lines of attack will work, not because they have developed a formal understanding of the properties of materials or contents of the problem, but because they have developed a feel for them, through experience" (p. 33). Anderson (1982), Fitts (1964), and Woolnough and Allsop (1985) have all emphasized that skill acquisition requires practice and experience.

Padilla, Okey and Dillashaw (1983) have demonstrated a high correlation between students' achievement of science process skills such as controlling variables, and the students' level of intellectual development. Many authors working in the Piagetian developmentalist paradigm (Good, 1977; Lawson, 1985) have argued that the ability to isolate and control variables in novel situations requires an abstracted and generalised reasoning pattern only available to those students who have attained formal operations.
Millar and Driver (1987) have agreed that process skills are learned in particular contexts, the skills are context dependent and are therefore difficult to transfer to new situations.

Rowell and Dawson (1989) proposed a theory of learning that integrates Piagetian learning theories, recent cognitive process psychology, and artificial intelligence (AI). Their integrated theory proposes that cognitive skills such as control of variables are learned in particular contexts and that these domain specific skills are stored in long term memory as domain specific schemata. McClelland and Rumelhart (1988) and Rowell and Dawson (1989) believed that schemata are in hierarchical order, that is, one schema may subsume other schemata. They defined a schema as a structure governing a body of inferences such that, when activated, it organizes comprehension of event-based situations. Chi and Rees (in Rowell and Dawson, 1989) suggested that "a group of production rules with the same goal element can be viewed as a schema" (p. 51).

When the learner has practised controlling variables in several different contexts and has developed several domain specific schemata for that skill, these low level schemata
may be subsumed into a higher order schemata. This process involves the abstraction of a generalizable skill from the domain specific schemata into an abstract higher order schemata typical of formal operations.

Many studies (Friedler, Nachmias & Songer, 1989; Friedler & Tamir, 1986; Johnson & Wham, 1982) have found that students' ability to plan and control variables improved when they participate in, and plan open-ended science investigations. Further, Friedler and Tamir (1986) and Johnson and Wham (1982) have argued that inquiry oriented, investigation style laboratory work is cognitively demanding as the students are involved in activities that force them to simultaneously apply functional knowledge of subject matter, proficiency in laboratory skills and intellectual enquiry skills. Tamir and Amir (1987) recommended that these skills should be taught rather than hoping they will just 'happen' as a result of laboratory experience. Hackling and Garnett (1991) suggested that some of the science process skills be initially taught in a non-laboratory situation so that the additional burdens from working with apparatus are avoided.
Assessment of Skills

Stannard (1982), Moriera (1980), and Bryce and Robertson (1985) have argued that while there has been a shift in emphasis of science courses toward a laboratory-centered approach, there has not been a change in the methods of evaluating students' achievement to accommodate the shift. Bryce and Robertson (1985) reported that there is evidence that many students can do science much better than traditional assessment permits them to demonstrate. Hofstein and Lunetta (1982) considered the area of practical skills assessment to be a neglected aspect of research.

Tamir (1972) supported the claim made by Kelly and Lister (1969) that practical work involves both manual and intellectual skills which are distinct from those used in non-practical work. These skills, Tamir argued, cannot be assessed by written examinations. He noted that the correlation between the assessment of these skills using written examinations compared to the result observed in a practical laboratory situation is very low. Similarly, Robinson (1969) found a very low correlation between laboratory practical examinations and written paper-and-pencil tests of science process skills. Lunetta et al., (1981) discussed four methods of evaluating laboratory
work. They are written laboratory reports, written test items, practical tests, and observational assessment.

Written Laboratory Reports

Written reports can be used to assess those skills associated with planning and design, analysis and interpretation, and application phases of science investigation.

Written Tests

Pencil and paper tests can be used to assess knowledge of techniques and principles underlying laboratory procedures, skills in planning and design, analysis, and application phases of laboratory activity. On the other hand, written tests cannot be used to assess apparatus skills associated with the performance phase of investigations.

Practical tests

Practical tests are used to assess acquisition of apparatus skills, observational skills, and more complex problem solving and science process skills. Practical tests have been a feature of many external examinations in England and Israel for a number of years. Their use has not been extensive because they are more difficult to design, construct, and administer than are written tests.
Observational Assessment

This system of assessment requires the teacher to observe and rate each student during normal laboratory activities. During this time, the teacher assesses behaviours in planning and design, manipulative skills, conduct of experiment, collection of data, responsibility, initiative, and work habits. This is not an easy method because the teacher is usually busy with classroom management tasks. There is not enough time to observe each individual student in a single period of practical work with a class of twenty-five to thirty students. Because of these difficulties this technique lacks reliability.

In order to obtain valid and reliable data about the range of inquiry skills that students possess, a combination of written tests and practical tests is favoured. They are complementary to each other. While practical tests can measure skills that cannot be assessed by written tests, written tests can measure some of the skills in a more efficient way.

Several pencil and paper tests of science process skills have been reported in the literature, three are described here. First is the Test of Integrated Process Skills (TIPS) produced by Okey and Dillashaw (1980). It is a 36-multiple choice item written test designed for grades 7-12 and takes
approximately 45 minutes to complete. It uses non-specific content to assess the acquisition of five science process skills: stating hypotheses, defining operationally, identifying variables, designing investigations, and graphically analyzing data. Second is the Test of Integrated Science Processes (TISP) developed by Tobin and Capie (1982). It is a 24-multiple choice item written test suitable for students from middle school to college level. It uses non-specific content material to assess 12 objectives for planning and conducting an investigation. Third is the Processes of Biological Investigation Test (PBIT) developed by Germann (1989). PBIT is a 35 item written test designed to measure high school biology students’ achievement of skills in developing hypotheses, making predictions, identifying assumptions, analyzing data, and formulating conclusions.

More recently, practical tests have been developed to assess process and apparatus skills. The test instruments for the Second International Science Study (SISS) reported by Jacobson and Doran (1988) used a combination of a content-based written test and a practical, science process laboratory skills test. It was prepared by the International Association for the Evaluation of Educational Achievement (IEA) for the Second IEA Science Study. In the practical, science process laboratory skills tests the
students were asked to design experiments, manipulate equipment, observe, record, and interpret data. Bryce (1989) reports the work of the Techniques for the Assessment of Practical Skills (TAPS) research team. The TAPS research team identified different kinds of skills required in practical science, and developed a practical test item bank for the assessment of science practical skills. The team performed three phases of research. In the first phase (1980 – 1983), they developed practical test items of three minutes duration to assess skills in six basic skill areas. These are observational skills, recording skills, measurement skills, manipulative skills, procedural skills and following instructions. These items can be found in Bryce, McCall, MacGregor, Robertson and Weston (1983). In the second phase (1983 – 1987), the TAPS research team developed structured, practical tasks of 30 minutes duration for the assessment of process skills in the areas of inference and selection of procedures (Bryce, McCall, MacGregor, Robertson & Weston, 1988). In the third phase (1987 – 1990), the team developed semi-structured investigations of three to four hours duration for the assessment of investigative skills in four areas. These are generative skills, experimental skills, evaluative skills and recording and reporting skills.
Hackling and Garnett (1990) developed a practical test instrument to assess high school students' attainment of skills associated with problem analysis and planning, collecting information, organizing and interpreting information, and concluding. In this practical test, the students were asked to verbalize their thoughts as they worked on the task and their performance was videotaped for analysis.

In the Second International Science Study (SISS), American students in Grades Five and Nine did well in manipulating scientific equipment, measuring, and recording data, but not in explaining, designing experiments, and reasoning (Jacobson and Doran, 1988). Tamir (1989) reported the results of the SISS science process laboratory tests of ninth grade (15-year-old) students in Israel. It was found that most of the Israeli students were comfortable with the skills of observing, planning, and reporting. However, they did not perform well in inferring, drawing conclusions and performing simple calculations. Reasoning skills like controlling variables and drawing conclusions are the common weaknesses in all these results. These data are consistent with the results of the study by Hackling and Garnett (1990) of Year 12 science students' attainment of science investigation skills in Western Australia. They
found that students used systematic measurement procedures, but were not competent in planning investigations and controlling variables.

The final study considered in this review is the Assessment of Performance Unit (APU) Science Project. It was part of a national monitoring exercise which looked at the performance of science students in England, Wales and Northern Ireland. They surveyed three age groups of students, 10-11 year old primary school students, 13 and 15 year old secondary school students. Written tests, individual practical tests and group practical tests were developed and used to assess students' performance in graphical and symbolic representation, use of apparatus and measuring instruments, observation, interpretation and application, planning of investigations, and performance of investigations (Murphy, 1988). The main findings of this research include the following. First, the content and the setting within which the problem or the investigation is presented affect the performance of the students. Factors like familiarity of the problem to be investigated, context in which the problem is presented, and experience of the students affect their performance. Second, it was found that students' ability to identify and control
variables depend on their conceptual understanding of the variables and their effects (Murphy, 1988).

In summary, many writers in science education now believe that the main benefit of laboratory work in high school science is the acquisition of skills that are related to the processes of science. It has been emphasized that acquisition of science investigation skills requires considerable practice and experience. It has been claimed that laboratory work involves skills which are distinct from those used in non-practical work and therefore practical skills are best assessed by a combination of written and practical tests.
CHAPTER 3

Methodology

Introduction

This chapter outlines the methodology of the study. The development of the two test instruments, the subjects of the study, data collection and analysis procedures are described in this chapter.

Instrument development

The development of the test instruments was guided by three factors. These were benchmark statements, a content validity grid, and formative evaluation data from pilot studies.

Benchmark Statements and Objectives

Prior to the development of the test instruments, benchmark statements were prepared to describe the range of science investigation skills and standards of performance that should be expected of Tongan Form 5 General Science students. These were prepared using the Tongan Form 5 General Science syllabus and the Year 10 Science Process Skills Benchmark Statements prepared for the Western
Australian Monitoring Standards in Education Project (Western Australian Ministry of Education, 1990). Behavioural objectives were prepared to describe the behaviours that should be observed in students who have achieved the benchmark standards. Benchmark statements and behavioural objectives were prepared for the following skill areas: problem analysis, planning, manipulating equipment, collecting information, organizing information, interpreting information, and communicating information. As an example, the benchmark statement and behavioural objectives for Problem Analysis are as follows:

**Benchmark**

The student demonstrates understanding of a problem by developing questions for investigation, proposing hypotheses for testing and predicting the outcome of specific investigations.

**Behavioural Objectives**

The student is able to:

(a) state the problem to be investigated;

(b) develop questions for investigation when background information is provided;

(c) identify the variables that could influence the phenomenon to be investigated;

(d) identify dependent and independent variables;
(e) develop hypotheses that state the relationship between the dependent and independent variables; and
(f) predict the outcome of particular investigations or events.

The rest of the benchmark statements and the behavioural objectives can be found in Appendix 1.

When the objectives had been prepared, it was realized that it would be impossible to test all of the objectives in this project. A representative number of objectives was then selected to be assessed and the assessment methods were decided. A written and a practical test were considered necessary in order to cover the range of science investigation skills selected for the study. Test items were either adapted from available sources (Dillashaw & Okey, 1980; Tobin & Capie, 1982) or constructed. In the adaptation of items from existing tests, the items were revised to fit the purposes of the study. In the construction of test items, the following points were considered:

- the questions are relevant to the Form 5 Science curriculum and the Tongan culture;
- the language and terms used in the questions can be understood by the subjects;
o the subjects use a specific skill in obtaining
the correct response for an item;
o the questions are clear and direct;
o the questions are free of misleading statements
or clues; and
o supplementary materials are easy to use.

These considerations were carried out mainly to avoid
the threats against the goodness of fit of the items to the
Rasch model. This is discussed further in Chapter 4.

Written Test of Science Process Skills

This test was designed to assess the objectives related
to problem analysis, planning, collecting information,
organizing information, interpreting information, and
communicating information. For each benchmark, test items
were either modified from existing test papers (Dillashaw &
Okey, 1980; Tobin & Capie, 1982) or created. A grid of
objectives versus test items was drawn to ensure coverage
of the objectives in the test, the final version of the
grid is presented in Appendix 2. This grid specifies the
number of questions that assess the attainment of the range
of science investigation skills described by each
benchmark. This ensured the content validity of the test.

Initially, a test of 42 items was produced. These items
were then divided into two groups to form two tests with 21
questions each. Each test took about 50 minutes to be
completed by average year 11 West Australian science students. Both tests were trialled on year 10 and year 11 science students in a large high school in the metropolitan area of Perth. It was assumed that the Tongan Form 5 (17 years of age) science students would be equivalent to the above-average Year 10 and the below-average Year 11 Perth science students. This assumption was based mainly on the results of a comparison of the science syllabi that are used in Perth and in Tonga. The results of the tests were analysed using ASCORE (Andrich, Sheridan & Lyne, 1991), a Rasch model computer program for analysing the characteristics of test instruments, to determine the goodness of fit of the items to the Rasch model. ASCORE is further discussed in the data analysis section. Item misfit is interpreted in test construction using the Rasch model as evidence of item inappropriateness. The misfitting items were inspected to reveal the causes for their deviation from the model, and depending on these causes, the misfit items were either modified or discarded. With the assistance of the content validity grid, new items were written to replace the deleted items. The test items were then validated against the objectives by three staff members and two postgraduate students in the Science Education Department of the Edith Cowan University. Their comments were considered and the items were once again revised.

The reading level of the test was analysed. This was checked using the Fry Readability Graph (Fry, 1977). This
procedure uses sentence complexity and word difficulty to
determine readability. The readability of a text book is
determined by plotting the average number of sentences and
syllables per 100 words on the Fry Readability Graph. The
average number of sentences per 100 words in the written
test of science process skills is 9.5, and the average
number of syllables per 100 words is 138.5. When these were
plotted on the Fry Readability Graph, the readability
estimate of the test was found to be at the upper fifth
grade level, that is, suitable for 10 year old students
which was considered satisfactory. The readability of the
questions that were between 40 and 99 words long was
determined using Fry’s readability formula for short
passages (Fry, 1990). This formula is based on word and
sentence difficulty. The readability estimate of the short
passages was found to be at the seventh grade level, that
is, suitable for 12 year old students.

The second draft test instrument of 27 items was then
trialled on year 10 and year 11 students in another high
school in Perth. The results were analysed using ASCORE.
The test was then revised for the last time and the final
version of the test was prepared. The final version of the
test includes 11 multiple choice questions and 14 open-ended questions, and is presented in Appendix 3. A marking
key for the test was developed together with the test. This
can be found in Appendix 4.

In Tongan schools, the language of instruction is
English. However, to ensure that Tongan students understood the test, two Form 3 (14 years old) and two Form 5 (16 years old) Tongan General Science students went through the test with the investigator and identified terms and questions they found difficult to understand. These difficult English language terms were noted and were explained to the subjects in each administration of the test in the main study. Further, the subjects were allowed to ask for definitions of terms in Tongan, and to write in Tongan if they could not express their ideas in English.

Practical Test of Science Apparatus Skills

This test was designed to assess students' attainment of skills of using scientific equipment to measure length, mass, liquid volume and temperature, perform simple chemical tests, and use a Bunsen burner to heat a test tube of liquid. A practical test was deemed necessary since written tests could not validly measure the skills involved in manipulation of scientific equipment (Bryce & Robertson, 1985; Kelly & Lister, 1969; Tamir, 1972). The objectives assessed in the practical test can be found in Appendix 5. Its development was similar to the procedure followed in the development of the written test of science process skills. However, special care was taken to ensure that the test items were efficient in terms of time and equipment available. Construction of test items followed the style
employed by Jacobson and Doran (1988). This study employed the stations method to administer the test to students. The activities were arranged at five stations around the room. The task at each station can be completed by an average Form 5 student in 10 minutes. The students were required to visit the stations and complete the activities. This method was selected because of its efficiency in terms of time and equipment.

Product and process evaluation methods were both used in the assessment of students' apparatus skills. Where students were required to make measurements of length, mass and liquid volume, and in the chemical tests activities, the products of the students' work were assessed. In the activities that required students to measure temperature and light a Bunsen burner, the processes of using the apparatus were observed by the investigator and assessed. Hence the investigator observed the students at one of the stations.

The first trial of the test was conducted with third year biology students in one of the tertiary institutions in Perth. This was not only to check the administration techniques and equipment needed for the experiments, but also to ascertain the highest level of performance that could be expected of high school science students. This data was important in the development of the marking key.
The second trial of the practical test was conducted with Year 10 and Year 11 students at a high school in Perth. During this trial, great care was taken to note any modification that needed to be made to the administration procedures. The following are examples of the changes made after the second trial of the test:

- instruct the students that the station numbers relate to the pages in the test booklet;
- seal the test tube and the jam jar in station 4 on the estimation of volume task to prevent students from using a measuring cylinder to measure the amounts of water in them; and
- check the stations after each student to see that in station 1, cursors on the triple-beam balance are pushed to the right end of the scale and there are no marks on the rulers.

The Test Instrument, Marking Key and the Administration Guide are presented in Appendices 6, 7, and 9 respectively.

Subjects

There are five main groups of islands in Tonga. They are Tongatapu, Vava’u, Ha’apai, ‘Eua, and Niua. There are 11 high schools in Tongatapu, five in Vava’u, three in
Ha'apai, one in 'Eua, and one in Niua. Form 5 General Science is taught in all of the high schools in Tonga. There is a total of about 700 Form 5 General Science students in Tonga and about three quarters of them live on the main island. The number of Form 3 and Form 5 General Science students in Tonga is roughly the same.

The written test of science process skills was administered to 206 Form 5 and 214 Form 3 General Science students, that is approximately 40 students (20 from each level) from each of the 10 schools selected from the 21 schools in Tonga. From the above 420 students, 100 of them undertook the practical test, 50 Form 5 and 50 Form 3 students. Below is a table giving the number of schools and students selected for the study.

<table>
<thead>
<tr>
<th>Island groups</th>
<th>Number of schools</th>
<th>Number of schools tested</th>
<th>Number of students tested</th>
<th>Written test</th>
<th>Practical test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tongatapu</td>
<td>11</td>
<td>6</td>
<td>260</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>Vava'u</td>
<td>5</td>
<td>2</td>
<td>80</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Ha'apai</td>
<td>3</td>
<td>1</td>
<td>40</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>'Eua</td>
<td>1</td>
<td>1</td>
<td>40</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>10</td>
<td>420</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>
The major criteria considered in selecting the schools for the study included the following: the equivalence of the Form 3 and Form 5 groups, the location of the schools, the socio-economic background of the students, and the past performances of the schools in the national and external examinations.

The equivalence of the Form 3 and Form 5 groups is necessary for the validity of the research method employed in this project. This is further discussed in the next section. However, to ensure the equivalence of the two groups, the schools selected for the study satisfied the following:

- the Form 5 General Science program is a continuation of that in Form 3;
- teachers in the two groups have similar background and experiences; and
- the ability levels of the students are the same.

The schools selected covered a wide range of socio-economic background. From the low socio-economic background of remote island life to that of the village life on the main island, and to the relatively higher socio-economic background of life in the town area. In terms of location, four schools were selected from the town area, and two from
the village areas of the main island. The other four schools were selected from the outer islands. In terms of their performances in the past national and external examinations, two top performing schools and two from the lowest performing schools were selected. The other six range from below average to above average.

In Tonga, there is one all-girls and two all-boys schools. These were all included in the study. Equal numbers of male and female students were included in the sample.

This study was conducted during the final weeks of the 1990 school year (mid-September to mid-October). Form 5 students had just completed the Tongan Form 5 General Science syllabus. Form 3 students represent the entry behaviour of students commencing the General Science program.

**Procedure**

This study tested Form 3 and Form 5 General Science students. These two sets of data were to be compared to identify the skills that students have attained after completing Form 4 and Form 5 General Science programs. Form 3 students were tested to identify the skills that students
possess as they enter Form 4 science. Form 5 students were tested to identify the skills that students possess as they leave Form 5 science. Both groups were administered the same test instruments. In this procedure, the assumptions about the equivalence of the groups are very important, because the procedure is valid only when these assumptions are true. The schools selected for the study had satisfied the conditions that would secure equivalence of the two groups.

The teachers and students were both assured that the results of the tests would be used only for the purposes of the study, and the individual students, teachers and schools would remain anonymous. This was important because since the schools were tested on different dates, students might try to score well by seeking information from the students that had been tested. To avoid this problem happening in individual schools, both groups (Form 5 and Form 3) at a school were tested on the same day, either together at one time or one after the other. Furthermore, to prevent the tests from reaching those schools yet to be tested, all the test papers were collected after the test.

Form 5 students required one hour to complete the written test of science process skills and 50 minutes to complete the practical test of science apparatus skills. To
compensate for the slower reading rate of the Form 3 students, they were given an extra 10 minutes for each test. They were given 70 minutes for the written test and 60 minutes to do the practical test. In each administration of the tests, five minutes was devoted to an explanation of the test instructions and giving definitions of difficult terms. Another five minutes was given as reading time. Setting up the stations in the practical test required at least 30 minutes, and setting up the room and materials required for the written test took an average of 15 minutes.

Data Analysis

Data collected throughout this study were used for two purposes. First, to determine the characteristics of the test instruments, and second, to assess the skill attainment of students. ASCORE was used to analyse the data to determine the characteristics of the test instruments. The performance of the students was analysed using the STATISTICAL ANALYSIS SYSTEM (1988) computer program.

ASCORE, a Rasch model computer program, was used to assess the psychometric properties of the measuring instruments. It was designed as a general program for analysing a range of data and is based on the Extended
Logistic Model (ELM), an elaboration of Rasch’s Simple Logistic Model (Rasch, 1960; 1980). Detailed discussion of the ELM is provided by Andrich (1982; 1985; 1988). The program can be used when the number of categories is two as for dichotomously scored multiple choice items that are scored as right (one mark) or wrong (zero marks), and also for open-ended items scored in multiple categories such as zero marks, one mark, two marks or three marks.

Throughout the instrument development process, ASCORE was used to analyse the results of the trials. The main purpose of these analyses was to determine the goodness of fit of the test items to the Rasch model. If the data fit the model, the data can be used to calibrate items or measure persons (Wright, 1977). The items that did not fit the model were analysed to identify the sources of deviation from the model. There are a number of identifiable sources of threat against the goodness of fit of a test to the Rasch model. These include: item heterogeneity, item bias, the extent to which the test is speeded, guessing, non-independence of responses, heterogeneous item discrimination, and heterogeneous person sensitivity (Gustafsson, 1979). Once the sources of deviation from the model were identified, the items were revised in such a way as to improve the goodness of fit of the items to the Rasch model. The raw scores obtained from
the administration of the test to the Tongan students were analysed using ASCORE to determine the suitability of the data for the calibration of the items and for the measurement of the Tongan subjects. The results of this analysis are discussed in the next chapter.

STATISTICAL ANALYSIS SYSTEM (SAS) is a computer program that has routines for describing data and generating statistical analyses. It was used to provide various descriptive statistics and to perform t-tests on the results of the two tests. The following tasks were performed using SAS.

(i) On both tests, means and standard deviations for all individual items, subtests, and total scores for all Form 5 students, all Form 3 students, Form 5 male students, and Form 5 female students were calculated.

(ii) On both tests, t-tests were used to compare the means on subtests and total scores for all Form 5 versus all Form 3 students, and Form 5 male against Form 5 female students.

The results of these analyses are presented in the Results Chapter.
Evaluation of the Test Instruments

Introduction

The written test of science process skills was developed through three pilot studies, the practical test of science apparatus skills was developed through two pilot studies. Data from pilot studies of the written test of science process skills and from both tests in the main study were analysed using ASCORE to evaluate the psychometric properties of the instruments. This chapter presents these formative evaluation data for the second trial of the written test and for the final version of both tests used in the main study.

The Written Test of Science Process Skills

Data from the Second Trial

The second trial of the written test was conducted with Year 10 and 11 students in a Perth high school. ASCORE was used to identify the items that had a poor fit with the Rasch model. These items were then analysed to reveal the cause of their misfit and were modified accordingly.

The results for the multiple choice section of the test were analysed first. The item-trait interaction overall
test-of-fit on the multiple choice part of the test yielded a Chi-square of 25.556, with nine degrees of freedom, and probability less than the acceptance level of 0.05. This result indicated that the multiple choice part of the test overall did not fit the model. The individual item tests-of-fit scores were used to identify the items that did not fit the model. Table 2 gives the results of the test-of-fit on individual items. It lists the items in order of fit to the model based on probability, commencing with the item of best fit.

<table>
<thead>
<tr>
<th>Item number</th>
<th>Chi-square</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>.001</td>
<td>.976</td>
</tr>
<tr>
<td>5</td>
<td>.012</td>
<td>.909</td>
</tr>
<tr>
<td>1</td>
<td>.394</td>
<td>.516</td>
</tr>
<tr>
<td>10</td>
<td>1.107</td>
<td>.271</td>
</tr>
<tr>
<td>8</td>
<td>1.515</td>
<td>.194</td>
</tr>
<tr>
<td>4</td>
<td>1.622</td>
<td>.178</td>
</tr>
<tr>
<td>6</td>
<td>1.752</td>
<td>.161</td>
</tr>
<tr>
<td>2</td>
<td>3.928</td>
<td>.018</td>
</tr>
<tr>
<td>3</td>
<td>4.721</td>
<td>.000</td>
</tr>
<tr>
<td>7</td>
<td>10.504</td>
<td>.000</td>
</tr>
</tbody>
</table>

Note. Degrees of freedom = 1

These data indicated that items 2, 3, and 7 caused the deviation of the multiple choice section of the test from the model. When only items 3 and 7 were removed from the
analysis, the multiple choice section of the test had a better fit with the Rasch model ($\chi^2 = 10.844 \ (7.0), \ p = 0.1193$). Items 2, 3 and 7 were analysed to reveal the causes of their deviation from the model. It was decided that item heterogeneity was the main cause. Therefore faults in the structure and wording of the items were identified and amended to increase homogeneity of these items. The original and modified versions of Item 3 are presented below as an example of how trial data was used to improve test items.

In the second trial, Item 3 was as follows:

3. Sue wanted to find out what might affect the growth of bean seedlings. She had ten identical beans. She placed a bean wrapped in moist tissue paper in each of ten identical test tubes. She put five of the tubes in a sunny window. She put the other five in a dark refrigerator. She measured the height of the bean seedlings in each group after one week.

Which of the following variables might cause differences in lengths of the bean seedlings in this experiment?

a. Temperature and moisture
b. Moisture and length of test tubes
c. Light and temperature
d. Light and amount of time

Item 3 appeared in the final version of the test as follows:

3. Jane wanted to find out what might affect the growth of bean plants. She had ten identical bean plants and all of them were kept moist throughout the experiment. She put five of the plants in a warm sunny window. She put the other five plants in a dark refrigerator. She measured the height of all plants after one week.
Which of the following variables might cause differences in growth of the plants in this experiment?

a. Temperature only  
b. Moisture and light  
c. Time only  
d. Temperature and light

The name Sue was changed to Jane because it is a more common name in Tonga than Sue. The term 'plants' replaced 'seedlings' because it was suspected that some students might not understand the term 'seedlings'. Third, the part '... wrapped in moist tissue paper ...' was removed so that the difference in the amount of light obtained by the two groups of plants was more evident to the students. Fourth, the part '... in this experiment' was underlined to remind the students not to use their previous experiences with bean experiments to answer the question. Finally, to reduce the length of the question, the information about the use of identical test tubes was removed. Such improvements were made on the items that did not fit the Rasch model.

In the analysis of the results for the whole test, all 10 of the multiple choice items were treated as one item with multiple category scoring and the open-ended items as separate items each with multiple category scoring (Andrich, 1985; 1988). The Item-trait interaction overall test of fit indicated that the test overall did not fit the
Rasch model ($\chi^2 = 64.487 \ (17.0), \ p < .05$). The individual item tests-of-fit identified the items that caused the overall misfit of the test to the model. These data are reported in Table 3.

**TABLE 3**

Fit Order of the Total Items in the
Second Trial of the Written Test
(n = 33)

<table>
<thead>
<tr>
<th>Item number</th>
<th>Chi-square</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>OP02</td>
<td>.036</td>
<td>.846</td>
</tr>
<tr>
<td>OP07</td>
<td>.057</td>
<td>.806</td>
</tr>
<tr>
<td>OP15</td>
<td>.079</td>
<td>.772</td>
</tr>
<tr>
<td>OP06</td>
<td>.301</td>
<td>.570</td>
</tr>
<tr>
<td>OP14</td>
<td>.652</td>
<td>.401</td>
</tr>
<tr>
<td>OP08</td>
<td>.659</td>
<td>.399</td>
</tr>
<tr>
<td>OP13</td>
<td>.747</td>
<td>.369</td>
</tr>
<tr>
<td>OP01</td>
<td>1.871</td>
<td>.146</td>
</tr>
<tr>
<td>OP11</td>
<td>2.297</td>
<td>.103</td>
</tr>
<tr>
<td>OP12</td>
<td>3.083</td>
<td>.051</td>
</tr>
<tr>
<td>OP09</td>
<td>3.984</td>
<td>.017</td>
</tr>
<tr>
<td>OP03</td>
<td>4.026</td>
<td>.015</td>
</tr>
<tr>
<td>OP10</td>
<td>4.767</td>
<td>.000</td>
</tr>
<tr>
<td>OP16</td>
<td>5.240</td>
<td>.000</td>
</tr>
<tr>
<td>OP04</td>
<td>6.977</td>
<td>.000</td>
</tr>
<tr>
<td>Di01</td>
<td>6.999</td>
<td>.000</td>
</tr>
<tr>
<td>OP05</td>
<td>7.791</td>
<td>.000</td>
</tr>
<tr>
<td>OP17</td>
<td>14.922</td>
<td>.000</td>
</tr>
</tbody>
</table>

**Note.** OP = Open ended items, Di = Combined multiple choice items.
Degrees of freedom = 1

It is evident from Table 3 that item OP17 was the main cause of the misfit of the test because of the big jump in the Chi-square value. However, these data show that items OP09, OP03, OP10, OP16, OP04, Di01, OP05, and OP17 did not fit the model well. Items OP04, OP05, and OP06 were
deleted. Item OP06 was deleted because it was a continuation of Item OP05. Item OP17 was not deleted but modified to maintain coverage of objectives. Item Di01 was the combined multiple choice items and the items responsible for the deviation of the multiple choice part were identified and revised as previously described. Items OP03, OP09, OP10, OP16, and OP17 were examined to reveal the causes of their misfit and were modified. The original and final version of item OP16 are presented below as an example of the improvements made to the misfit items.

In the second trial, item OP16 was as follows:

16. A scientist performed the following experiment to test an hypothesis:

The hypothesis was: Left-handed people play better tennis than right-handed people.

The experiment:

A list of all the tennis players in the country was compiled. The scientist then randomly selected from the list two left handed (one male and one female) and two right handed (one male and one female) players. They were allowed to play against their own sex. Each match was won by a left handed player.

Question:

Write down all the faults in the experimental design that you can think of.
In the final version of the written test, item OP16 became item OP13 and it is as follows:

13. A scientist performed the following experiment to test an hypothesis:

The hypothesis was: *Left-handed people play better tennis than right-handed people.*

The experiment:

A list of all the tennis players in the country was compiled. The scientist then randomly selected from the list one left handed and one right handed player. A match between the two players was won by the left handed player.

Question:

Write down all the faults in the experimental design that you can think of.

The item was modified so that the faults relating to sample size and control of variables became more evident. The final version of the written test was composed of 11 multiple choice and 14 open ended items.

Data from the Main Study

Since the tests were designed mainly for the Form 5 students, the date collected from the Form 3 students during the main study in Tonga were not included in these analyses. The data referred to in this section are those that were collected from the Form 5 students in Tonga during the main study. These data were analysed using AS CORE. The multiple choice section of the written test was analysed first. The item-trait interaction overall test of fit for the multiple choice section of the test indicated
an overall reasonable fit to the model
\( \chi^2 = 45.655 \ (30.0), \ p = 0.0335 \). The individual item
tests-of-fit are reported in Table 4.

**TABLE 4**

Fit Order of the Multiple Choice Items in
the Written Test of Science Process Skills
\( (n = 201) \)

<table>
<thead>
<tr>
<th>Item number</th>
<th>Chi-square</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (1)a</td>
<td>1.287</td>
<td>.724</td>
</tr>
<tr>
<td>8 (8)</td>
<td>2.173</td>
<td>.523</td>
</tr>
<tr>
<td>6 (6)</td>
<td>2.217</td>
<td>.514</td>
</tr>
<tr>
<td>5 (5)</td>
<td>2.280</td>
<td>.502</td>
</tr>
<tr>
<td>10 (10)</td>
<td>2.400</td>
<td>.478</td>
</tr>
<tr>
<td>4 (4)</td>
<td>2.785</td>
<td>.408</td>
</tr>
<tr>
<td>9 (9)</td>
<td>4.852</td>
<td>.158</td>
</tr>
<tr>
<td>3 (3)</td>
<td>4.961</td>
<td>.149</td>
</tr>
<tr>
<td>11 (-)</td>
<td>6.159</td>
<td>.077</td>
</tr>
<tr>
<td>2 (2)</td>
<td>7.497</td>
<td>.029</td>
</tr>
<tr>
<td>7 (7)</td>
<td>9.045</td>
<td>.000</td>
</tr>
</tbody>
</table>

*Note. aNumber of the same item as it
appeared on the second trial of the test.
Degrees of freedom = 3.00.*

Item 7 was the only one that did not fit the model well.
However, since the overall test of fit was adequate for an
exploratory study, item 7 was not rejected.

The multiple choice section was then treated as one item
with multiple category scoring, and the results for the
whole test were analysed. The item-trait interaction
overall test of fit indicated that the test overall does
not fit the Rasch model \( \chi^2 = 579.745 \ (42.0), \ p \ll .05 \).
The individual tests-of-fit were examined to identify the
items that were responsible for the misfit of the test. The
results of these tests are reported in Table 5.
It is apparent from Table 5 that too many items did not fit the model. A closer inspection of the analysis of the misfitting items revealed that many items were not discriminating well, some were overdiscriminating while others were underdiscriminating. In some items, some of the lower performing students provided the correct answer while some of the higher performing students provided wrong responses. After a thorough consideration of the sources of threat against the fit of a test to the Rasch model, it was decided that there were three possible causes for the deviation of the test from the Rasch model. First is item heterogeneity, that is the items were measuring different
abilities, which is a violation of the assumption of unidimensionality in the Rasch model. This test measured the attainment of skills in six different process skill areas. Items in each skill area required the students to apply different cognitive skills from those required in the other areas. Testing all science process skill areas in one test and combining the scores as a measure of process skill attainment therefore needs reconsideration. The alternative is to test each skill area separately. In view of this concern, the items have been grouped into subtests directed at separate skill areas for the analysis of student performance described in the next chapter. Another source of item heterogeneity is the mix of multiple choice and open-ended items. Different cognitive processes are used in responding to these different types of test items. Further development of this test should consider the selection of items requiring a single response style. The other reason for the deviation of the test from the model is that the scores were very low. It is evident from the results of the written test that most of the Tongan Form 5 General Science population were well below the standards described in the benchmark statements. Therefore this test missed its target because the population performed at a level below the expected standards. This creates a bias in the analysis of fit of the test to the Rasch model.
Practical Test of Science Apparatus Skills

In the analysis of the results from the practical test, the item-trait interaction overall test of fit for the whole test indicated a very good fit to the Rasch model \( \chi^2 = 82.006 \) (78.0), \( p = 0.3562 \). The individual tests-of-fit, reported in Table 6, confirmed that all the items fit the model well.

**TABLE 6**

<table>
<thead>
<tr>
<th>Item number</th>
<th>Chi-square</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>.633</td>
<td>.966</td>
</tr>
<tr>
<td>9</td>
<td>2.569</td>
<td>.857</td>
</tr>
<tr>
<td>10</td>
<td>2.634</td>
<td>.849</td>
</tr>
<tr>
<td>11</td>
<td>3.097</td>
<td>.791</td>
</tr>
<tr>
<td>13</td>
<td>3.138</td>
<td>.786</td>
</tr>
<tr>
<td>5</td>
<td>3.250</td>
<td>.771</td>
</tr>
<tr>
<td>4</td>
<td>3.718</td>
<td>.707</td>
</tr>
<tr>
<td>14</td>
<td>4.560</td>
<td>.591</td>
</tr>
<tr>
<td>2</td>
<td>5.695</td>
<td>.444</td>
</tr>
<tr>
<td>7</td>
<td>7.526</td>
<td>.256</td>
</tr>
<tr>
<td>8</td>
<td>9.559</td>
<td>.122</td>
</tr>
<tr>
<td>3</td>
<td>9.933</td>
<td>.104</td>
</tr>
<tr>
<td>12</td>
<td>12.266</td>
<td>.031</td>
</tr>
<tr>
<td>1</td>
<td>13.430</td>
<td>.011</td>
</tr>
</tbody>
</table>

*Note. Degrees of freedom = 6.00*

Although a small calibration sample will improve the chance of a good fit to the model, this result is highly satisfactory.

This chapter presented the results of the evaluation of
the test instruments. The results for the practical test are satisfactory because the items appeared to be unidimensional and fit the Rasch model. On the other hand, the results for the written test raised questions relating to the validity of preparing a single test that assesses all science process skills. This issue is addressed in the Conclusions and Implications Chapter.
CHAPTER 5

Results

Introduction

This chapter reports data regarding students’ performance on the tests. The report has five sections. First, the performance of students on the written test of science process skills is presented. Second, the performance of students on the practical test of science apparatus skills is reported. Third, gender differences in the performance of Form 5 students on the two tests is discussed. Fourth, the performance of Form 5 students in relation to the benchmark standards, and finally a summary of the main findings is presented.

Performance of Students on the Written Test of Science Process Skills

The performance of Form 3 and Form 5 students on each item of the written test is reported in Table 7. Form 3 students’ mean score for each multiple choice item is less than half marks (0.50), with the exception of MC01 and MC03. Question MC01 required the students to identify a testable hypothesis when the description of an investigation was given. In question MC03, a problem with a dependent variable was given and the students were required to identify variables that might affect the dependent variable. The two lowest means for Form 3 on the multiple
CHAPTER 5

choice questions were those for questions MC02 and MC06. These questions required the students to identify controlled variables when descriptions of investigations were given.

TABLE 7

Maximum Possible Score and Mean Scores
for Form 3 and Form 5 Students on Each Item of the Written Test

<table>
<thead>
<tr>
<th>Item number</th>
<th>Maximum possible score</th>
<th>Mean score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Form 3 (n = 214)</td>
<td>Form 5 (n = 206)</td>
<td></td>
</tr>
<tr>
<td>MC 01</td>
<td>1.00</td>
<td>0.57</td>
</tr>
<tr>
<td>MC 02</td>
<td>1.00</td>
<td>0.31</td>
</tr>
<tr>
<td>MC 03</td>
<td>1.00</td>
<td>0.56</td>
</tr>
<tr>
<td>MC 04</td>
<td>1.00</td>
<td>0.37</td>
</tr>
<tr>
<td>MC 05</td>
<td>1.00</td>
<td>0.42</td>
</tr>
<tr>
<td>MC 06</td>
<td>1.00</td>
<td>0.30</td>
</tr>
<tr>
<td>MC 07</td>
<td>1.00</td>
<td>0.38</td>
</tr>
<tr>
<td>MC 08</td>
<td>1.00</td>
<td>0.49</td>
</tr>
<tr>
<td>MC 09</td>
<td>1.00</td>
<td>0.32</td>
</tr>
<tr>
<td>MC 10</td>
<td>1.00</td>
<td>0.36</td>
</tr>
<tr>
<td>MC 11</td>
<td>1.00</td>
<td>0.40</td>
</tr>
<tr>
<td>OE 01</td>
<td>3.00</td>
<td>1.24</td>
</tr>
<tr>
<td>OE 02</td>
<td>2.00</td>
<td>0.19</td>
</tr>
<tr>
<td>OE 03</td>
<td>7.00</td>
<td>1.25</td>
</tr>
<tr>
<td>OE 04</td>
<td>3.00</td>
<td>0.15</td>
</tr>
<tr>
<td>OE 05</td>
<td>2.00</td>
<td>0.14</td>
</tr>
<tr>
<td>OE 06</td>
<td>7.00</td>
<td>0.84</td>
</tr>
<tr>
<td>OE 07</td>
<td>2.00</td>
<td>0.35</td>
</tr>
<tr>
<td>OE 08</td>
<td>5.00</td>
<td>1.37</td>
</tr>
<tr>
<td>OE 09</td>
<td>4.00</td>
<td>0.25</td>
</tr>
<tr>
<td>OE 10</td>
<td>4.00</td>
<td>0.26</td>
</tr>
<tr>
<td>OE 11</td>
<td>2.00</td>
<td>0.02</td>
</tr>
<tr>
<td>OE 12</td>
<td>3.00</td>
<td>0.02</td>
</tr>
<tr>
<td>OE 13</td>
<td>4.00</td>
<td>0.09</td>
</tr>
<tr>
<td>OE 14</td>
<td>4.00</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Note. MC = Multiple choice questions.
OE = Open ended questions.

For Form 5 students, the mean values were also low. The two highest and the two lowest means for the multiple choice questions were the same items on which Form 3
students had their highest and lowest means. The highest were those for questions MC01 and MC03, and the lowest were those for questions MC02 and MC06.

Form 3 students' mean score for each of the open ended items was below half marks. It is interesting to note that none of the Form 3 students achieved more than 25% of the possible score in question OE13. This question required the students to identify faults in the design of an experiment that had inadequate sample size and control of variables.

The highest two means were those for questions OE01 and OE08. Question OE01 required the students to identify where they could find information in a book when the contents page was given. In question OE08, students were given pieces of two different fruits and were asked to list five ways in which the cut surfaces were different. The lowest two means were those for questions OE11 and OE12. In question OE11, students were given a plan of an investigation and were then asked to identify the dependent variable. From the same plan, question OE12 required the students to identify additional variables that should have been controlled in the investigation, and to give reasons for their answer.

For Form 5 students, the only items for which the mean was above 50% were OE01 and OE07. Question OE01 has been described above. In question OE07, the students were provided with a problem and a mathematical expression. They were then asked to perform numeric calculations using the
CHAPTER 5

information given. The lowest two means as with the Form 3 students were those for questions OE11 and OE12.

In Table 8, scores are reported for each skill area or subtest included in the written test.

TABLE 8

Maximum Possible Score, Mean and Standard Deviation for Form 3 and Form 5 Students on the Skill Areas of the Written Test

<table>
<thead>
<tr>
<th>Skill area</th>
<th>Maximum possible score</th>
<th>Form 3 (n = 214)</th>
<th>Form 5 (n = 206)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>Planning and problem analysis</td>
<td>14.00</td>
<td>3.10</td>
<td>1.59</td>
</tr>
<tr>
<td>Collecting information</td>
<td>16.00</td>
<td>2.46</td>
<td>1.68</td>
</tr>
<tr>
<td>Organizing information</td>
<td>12.00</td>
<td>2.29</td>
<td>1.59</td>
</tr>
<tr>
<td>Interpreting information</td>
<td>14.00</td>
<td>2.28</td>
<td>1.67</td>
</tr>
<tr>
<td>Communicating information</td>
<td>7.00</td>
<td>1.50</td>
<td>1.18</td>
</tr>
</tbody>
</table>

Note.
* p < .01 two-tailed t-test for independent samples.

It is important to mention here that some students from both levels scored very well on the test. However, it should also be noted that although very high marks were attained by some students in each skill area, the mean values were very low. In fact no mean value was above 50% of the possible score for each skill area. Form 3 performed
best on planning and problem analysis, and worst on collecting information. Form 5 students performed best on communicating information, and worst on planning and problem analysis.

A two-tailed t-test for independent samples was used to compare the means of Form 3 and Form 5 students. It was found that for each skill area, there was a significant difference between the mean for Form 5 and the mean for Form 3 students. For example, on planning and problem analysis, the Form 5 mean score (4.28) is significantly different from the Form 3 mean score (3.10), $t(369) = 6.18$, $p<.01$.

Performance of Students on the Practical Test of Science Apparatus Skills

Students were assessed on their ability to use laboratory apparatus to carry out simple tasks such as using a Bunsen burner, making measurements and performing chemical tests. Table 10 reports the performance of Form 3 and Form 5 students on each item of the practical test.
Although some students attained high marks on each item, the mean score for each item was very low. The only mean score for Form 3 that was above 50% of the possible mark was that of item 2(c). For Form 3, the highest two means were those for items 2(c) and 4(b). Item 2(c) required the students to follow an instruction to carry out the iodine test for the presence of starch, and then to record their observations. Students were assessed on the observations they recorded. Item 4(b) instructed students to use measuring cylinders to measure out required volumes of water. This task tested not only their ability to read a volume of a liquid placed in a measuring cylinder, but also their ability to transfer a required volume of liquid from the measuring cylinder to another container. Assessment of this item was made on the amount of water they transferred.
The two lowest means for Form 3 were those for items 3(a) and 5(a). Item 3(a) required the students to light a Bunsen burner. The assessment was based on how they handled the burner, and the type of flame they obtained. Almost 80% of the students did not attempt to change the black-smoke yellow flame they obtained when using the burner to heat a test tube. In item 5(a), the students were asked to plan how they would investigate the presence of acid, base or water in four unknown liquids, using red and blue litmus paper.

For Form 5 students, the two highest means were those for items 2(c) and 2(d), which involved tests for starch using iodine. Item 2(c) required the students to record observations and item 2(d) required the students to draw conclusions based on their observations of colour changes. The two lowest means were those for items 1(c) and 4(a). Item 1(c) instructed the students to use a 200 g spring balance to measure the mass of a steel bar. The assessment was based on the accuracy of the value they recorded. Item 4(a) asked students to estimate volumes of water in two different containers. The values they recorded were assessed in terms of their accuracy. It is interesting to note that for items 1(c), 4(a) and 4(b), the means for Form 3 were higher than those for Form 5. Items 1(c) and 4(a) have been described above. Item 4(b) required the students to use measuring cylinders to measure required volumes of water.
Scores for the practical test are reported in Table 11 according to the skill areas included in the test. In each skill area, the mean value for Form 3 is always below half marks.

### TABLE 11

<table>
<thead>
<tr>
<th>Skill area</th>
<th>Maximum possible score</th>
<th>Form 3 (n = 50)</th>
<th>Form 5 (n = 50)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>Measurement of length</td>
<td>4.00</td>
<td>1.46</td>
<td>1.46</td>
</tr>
<tr>
<td>Measurement of mass</td>
<td>4.00</td>
<td>0.28</td>
<td>0.61</td>
</tr>
<tr>
<td>Starch test</td>
<td>8.00</td>
<td>2.92</td>
<td>2.14</td>
</tr>
<tr>
<td>Bunsen burner</td>
<td>5.00</td>
<td>1.30</td>
<td>0.95</td>
</tr>
<tr>
<td>Measurement of temperature</td>
<td>6.00</td>
<td>1.78</td>
<td>1.04</td>
</tr>
<tr>
<td>Estimation of volume of liquid</td>
<td>4.00</td>
<td>0.48</td>
<td>0.79</td>
</tr>
<tr>
<td>Measurement of volume of liquid</td>
<td>4.00</td>
<td>1.94</td>
<td>1.17</td>
</tr>
<tr>
<td>pH test</td>
<td>9.00</td>
<td>1.10</td>
<td>1.71</td>
</tr>
</tbody>
</table>

**Note.**

*p<.05, **p<.01 two-tailed t-test for independent samples.

For Form 3, the best mean was for the measurement of the volume of liquid. The lowest mean was for the measurement of mass. The means for the estimation of volume of liquid and for the pH test were also very low. For Form 5, the
highest mean was that for the starch test. The lowest mean was for the estimation of volume of liquid. The mean for the measurement of mass was also surprisingly low.

A two-tailed t-test for independent samples was used to compare Form 3 and Form 5 means. It was found that for each skill area, the Form 5 and Form 3 means were significantly different, with the exception of the estimation and measurement of volume of liquid.

Mean total test scores for Form 3 and Form 5 students on the practical test are reported in Table 12.

<table>
<thead>
<tr>
<th>Maximum Possible Score</th>
<th>Form 3 (n = 50)</th>
<th>Form 5 (n = 50)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>44</td>
<td>11.26</td>
<td>5.60</td>
</tr>
</tbody>
</table>

*Note. p<.01 two tailed t-test for independent samples.

The mean total scores for the test were very low. A two-tailed t-test for independent samples performed on the means showed that the Form 5 mean score (21.30) is significantly different from the Form 3 mean score (11.26), t(98) = 7.72, p<0.01.
Gender Differences in Performance of Form 5 Students on the Tests

Written Test of Process Skills

The performance of Form 5 boys and girls on the written test is reported in Table 13.

**TABLE 13**

Mean Scores for Form 5 Boys and Girls on Each Item of the Written Test

<table>
<thead>
<tr>
<th>Item number</th>
<th>Mean Boys (n = 105)</th>
<th>Mean Girls (n = 101)</th>
<th>Boys - Girls</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC01</td>
<td>0.75</td>
<td>0.73</td>
<td>0.02</td>
</tr>
<tr>
<td>MC02</td>
<td>0.37</td>
<td>0.33</td>
<td>0.04</td>
</tr>
<tr>
<td>MC03</td>
<td>0.62</td>
<td>0.63</td>
<td>-0.01</td>
</tr>
<tr>
<td>MC04</td>
<td>0.45</td>
<td>0.47</td>
<td>-0.02</td>
</tr>
<tr>
<td>MC05</td>
<td>0.50</td>
<td>0.46</td>
<td>0.04</td>
</tr>
<tr>
<td>MC06</td>
<td>0.40</td>
<td>0.44</td>
<td>-0.04</td>
</tr>
<tr>
<td>MC07</td>
<td>0.48</td>
<td>0.41</td>
<td>0.07</td>
</tr>
<tr>
<td>MC08</td>
<td>0.58</td>
<td>0.44</td>
<td>0.14</td>
</tr>
<tr>
<td>MC09</td>
<td>0.52</td>
<td>0.57</td>
<td>-0.05</td>
</tr>
<tr>
<td>MC10</td>
<td>0.58</td>
<td>0.62</td>
<td>-0.04</td>
</tr>
<tr>
<td>MC11</td>
<td>0.52</td>
<td>0.60</td>
<td>-0.08</td>
</tr>
<tr>
<td>OE01</td>
<td>1.84</td>
<td>1.79</td>
<td>0.05</td>
</tr>
<tr>
<td>OE02</td>
<td>1.01</td>
<td>0.30</td>
<td>0.71</td>
</tr>
<tr>
<td>OE03</td>
<td>2.68</td>
<td>2.56</td>
<td>0.12</td>
</tr>
<tr>
<td>OE04</td>
<td>0.77</td>
<td>0.60</td>
<td>0.17</td>
</tr>
<tr>
<td>OE05</td>
<td>0.42</td>
<td>0.28</td>
<td>0.14</td>
</tr>
<tr>
<td>OE06</td>
<td>2.98</td>
<td>3.11</td>
<td>-0.13</td>
</tr>
<tr>
<td>OE07</td>
<td>1.17</td>
<td>0.94</td>
<td>0.23</td>
</tr>
<tr>
<td>OE08</td>
<td>1.90</td>
<td>1.96</td>
<td>-0.06</td>
</tr>
<tr>
<td>OE09</td>
<td>0.65</td>
<td>0.64</td>
<td>0.01</td>
</tr>
<tr>
<td>OE10</td>
<td>1.21</td>
<td>1.45</td>
<td>-0.24</td>
</tr>
<tr>
<td>OE11</td>
<td>0.16</td>
<td>0.06</td>
<td>0.10</td>
</tr>
<tr>
<td>OE12</td>
<td>0.35</td>
<td>0.25</td>
<td>0.10</td>
</tr>
<tr>
<td>OE13</td>
<td>0.69</td>
<td>0.62</td>
<td>0.07</td>
</tr>
<tr>
<td>OE14</td>
<td>1.82</td>
<td>1.69</td>
<td>0.13</td>
</tr>
</tbody>
</table>

The third column in Table 7 gives the difference between the mean for the boys and the mean for the girls in each
item. A negative value indicates that the girls’ mean is higher than the boys’.

The two multiple choice questions on which the boys performed best in comparison to the girls were items MC07 and MC08. In item MC07, the students were asked to select the best plan for an experiment to test an hypothesis. Item MC08 tested the ability of the students to make simple extrapolations from a graph. The girls did best in comparison with the performance of boys on item MC11. This item required the students to calculate an unknown using information provided.

The two open-ended questions on which boys performed best in comparison with the girls were items OE02 and OE07. Item OE07 provided the students with a problem and a mathematical expression. They were then asked to carry out calculations using the information. Item OE02 required students to convert values in different units to the same unit and to perform a calculation. These two items tested similar skills. The girls did best in comparison with boys’ on items OE06 and OE10. Item OE06 required students to identify the controlled variable when a description of an investigation was given. Item OE10 tested the ability of the students to extract important information from a text.

A comparison of the performance of boys and girls on each skill area of the written test of science process skills is presented in Table 14.
TABLE 14
Mean and Standard Deviation for Form 5 Boys and Girls on Each Skill Area in the Written Test

<table>
<thead>
<tr>
<th>Skill area</th>
<th>Boys</th>
<th></th>
<th>Girls</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Standard deviation</td>
<td>Mean Standard deviation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planning and problem analysis</td>
<td>4.51 2.31</td>
<td>4.04 2.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collecting information</td>
<td>5.53 2.45</td>
<td>5.72 2.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organizing information</td>
<td>5.44 2.45</td>
<td>4.24 2.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interpreting information</td>
<td>4.90 2.17</td>
<td>4.71 2.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communicating information</td>
<td>3.05 1.61</td>
<td>3.24 1.54</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A two-tailed t-test of independent samples was performed on each pair of means. It was found that there were no significant differences between the performance of boys and girls in any of the skill areas.

Practical Test of Apparatus skills

A comparison of the performances of boys and girls on the skill areas of the practical test is reported in Table 15.
Once again, two-tailed t-tests for independent samples were carried out on the pairs of means to determine whether the differences between the means were significant. It was found that there were no significant differences between the means in any skill area.
CHAPTER 5

Performance of Form 5 Students in Relation to Benchmark Standards

This section reports the distribution of Form 5 students in relation to the benchmark standards for each skill area assessed by the written and practical tests.

The benchmark standards and behavioural objectives were developed after a close consideration of the situation of the Tongan students, the Form 5 General Science curriculum materials and the current views of various writers in science education regarding the standards and relevance of science investigation skills for secondary science students. The natures of the test instruments and their administrations were also considered. After these deliberations, the lines that separate the students into below, at, and above benchmark were drawn. Students scoring less than 40% of the possible mark were deemed to be below the benchmark standard. Those that scored between 40-55% of the possible mark were considered to be at the benchmark, and the students that scored higher than 55% of the possible mark were considered to be above the benchmark.

Written Test of Process Skills

Two hundred and six Form 5 students completed the written test of science process skills. The percentage of students below, at and above the benchmark standard for each skill area tested is given in Table 16.
It is interesting to note that for each skill area, most students performed below the benchmark standard. The students performed worst in the area of planning and problem analysis. In this skill area, 76% of the students were below the benchmark standard. The students performed best in the area of communicating information. In this skill area, 61% of the Form 5 students were at or above the benchmark standard.

**Practical Test of Science Apparatus Skills**

Only 50 of the 206 Form 5 students who had taken the written test of science process skills participated in the practical test. The percentage of students below, at and above the benchmark standard are reported in Table 17.
### TABLE 17
Percentage of Form 5 Students Below, At, and Above Benchmark Standards in the Skill Areas of the Practical Test

<table>
<thead>
<tr>
<th>Skill area</th>
<th>Below benchmark</th>
<th>At benchmark</th>
<th>Above benchmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurement of length</td>
<td>26</td>
<td>36</td>
<td>38</td>
</tr>
<tr>
<td>Measurement of mass</td>
<td>76</td>
<td>20</td>
<td>4</td>
</tr>
<tr>
<td>Starch test</td>
<td>6</td>
<td>16</td>
<td>78</td>
</tr>
<tr>
<td>Bunsen burner</td>
<td>26</td>
<td>54</td>
<td>20</td>
</tr>
<tr>
<td>Measurement of temperature</td>
<td>40</td>
<td>18</td>
<td>42</td>
</tr>
<tr>
<td>Estimation of volume of liquid</td>
<td>88</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>Measurement of volume of liquid</td>
<td>42</td>
<td>24</td>
<td>34</td>
</tr>
<tr>
<td>pH test</td>
<td>24</td>
<td>24</td>
<td>52</td>
</tr>
</tbody>
</table>

The two skill areas in which the students performed worst were the measurement of mass and the estimation of liquid volume. For the measurement of mass, 76% of the students were below the benchmark. For the estimation of liquid volume, 88% of the students were below the benchmark. The students performed best in the starch test and the pH test. In the starch test, 94% of the students were at or above the benchmark and in the pH test, 76% of the students were at or above the benchmark.
Summary of the Main Findings

The overall performance of the Tongan students on the two tests was very low. In the written test, the Form 5 mean score for each skill area was always below half marks. In relation to their performance on the other skill areas of the written test, the Form 5 students performed best in collecting and communicating information. They performed worst in the skill areas of planning and problem analysis. In the practical test, the total mean score for the Form 5 students was below half marks. However, in two of the eight skill areas tested, the Form 5 mean scores were above half marks.

A two-tailed t-test for independent samples was used to compare the means for Form 5 and Form 3 students in each of the skill areas tested in the written test. It was found that the Form 5 mean score was significantly different from the Form 3 mean score in each skill area. In the practical test, it was found that the Form 5 total mean score was significantly different from the Form 3 total mean score. In the skill areas tested in the practical test, the Form 5 mean score was significantly different from the Form 3 mean score in six of the eight skill areas.

The same test was used to compare the performances of Form 5 male and female students on the two tests. It was found that there were no significant differences between their performances in both tests. Finally, the Form 5 students were categorized into three divisions based on
their tests results. It was observed that in all skill areas of the written test, most students were still in the area of working towards the benchmark. However, in the practical test, it was revealed that in four of the eight skill areas tested, more students were above the benchmark standard than below.
Discussion

Introduction

This chapter interprets the results of the study. Since the tests were designed mainly for the Form 5 students, the discussion will be confined primarily to the Form 5 results. Form 3 students were tested to identify the science investigation skills that the students possess when they enter Form 4 science. The results for Form 3 will be compared with those for Form 5 to evaluate the effectiveness of the Form 4 and 5 science curriculum in developing students’ science investigation skills.

The performance of Form 5 students on the written test of science process skills is discussed first. Next is their performance on the practical test of science apparatus skills. The students’ performance on the two tests is analysed to reveal the weaknesses and strengths of the present Form 4 and Form 5 General Science curriculum with regard to students’ attainment of science investigation skills. An evaluation of gender differences in the performance of Form 5 boys and girls will follow, and finally, a discussion of students’ performance in relation to the benchmark standards will be presented.
Since the results of the Rasch analysis of the written test of science process skills showed that the test items did not hold together and fit the Rasch model, the interpretation of the results will attend mainly to the results for the individual items and the subtests for each of the skill areas. It is believed that the total scores for the written test do not have much meaning as the test is not unidimensional.

**Performance of Form 5 Students on the Written Test of Science Process Skills**

Informal discussions with Tongan secondary science teachers revealed that most of them were not really sure of the meaning of science process skills. They conceded that practical work had mainly been employed to reinforce and confirm the scientific theories and laws taught during science lessons. For them, the first priority had been to complete the content of the syllabus. Some of them indicated that one of the main reasons that prompted them to do practical work was to complete the requirements of the Ministry of Education. It is little wonder that in each of the skill areas tested, the mean for Form 5 students was always below half marks.

**Planning and Problem Analysis**

This area includes the skills in developing questions to
be investigated, formulating hypothesis for testing, identifying variables to be controlled and planning investigations. As might be expected, the students did worst in this area. During the trial session of the test instruments in Tonga, it was found that the terms controlled variables, hypothesis, dependent variables, and independent variables were totally new to most of the students. The definition of these terms were then given to the students at the beginning of each administration of the test. Further, each of the test items that referred to different categories of variables included an explanatory phrase in brackets after each use of the term. Tongan science teachers confirmed that students had very little experience in designing experiments. Hackling and Garnett's study (1990) of science investigation skill attainment by Year 12 science students in Western Australia, revealed that none of the students interviewed used the terms variable, independent variable, dependent variable, or control of variables, while working on an investigation. Planning and problem analysis are skills which are cognitively demanding, but unfortunately Tongan students have very little practice in them. Analyses of the implemented curriculum in the United States (Tamir & Lunetta, 1981), Israel (Friedler & Tamir, 1984) and Australia (Tobin, 1986) indicate that most high school laboratory work is of the recipe style that provides little opportunity for student planning. It is during problem analysis and planning that students learn to identify different types of variables.
The worst-answered question in the test was item OE11. It presented the students with a plan of an experiment and then required them to identify the dependent variable. Only one student gave the correct answer. This result is consistent with Tongan science teachers' reports that they had never discussed identification of variables, control of variables, and development of hypotheses with Form 5 students. Another question that required students to identify the dependent variable was item MC04. About 45% of the students chose the correct response. Higher performance observed in this item compared to the OE11 could be due to the difference in the level of difficulty and the structures of the two questions.

Collecting Information

This area includes skills of observing, describing and measuring specific characteristics and changes and recording of data. In comparison to their performance in other skill areas, the Form 5 students did very well in this area. Tamir (1989) reported that Israeli students were good at observation and reporting. Jacobson and Doran (1988) also reported that American students performed well in data recording. Lunetta and Tamir (1979) found that in many laboratory activities, students were required mainly to observe, measure and record results. Students would therefore have plenty of opportunity to practise and develop these skills.
Although the Tongan students performed best in the areas of collecting and communicating information, the level of their performance was very low. On the task of constructing a table, 50% of the students did not include units for the values recorded. Of those who included units in the table, only 30% put them at the top of the columns. On the observation task, 60% of the students reported four or more observed differences between the two fruits. However, only 40% of that proportion recorded which fruit had each of the observed characteristics. This reflects adequate experience in observation but little in recording of observations. On the task of drawing a diagram, 55% of the students did draw the correct structures observed but only three out of two hundred and six students included a title of the diagram. No one included a scale, and most of them either labelled incorrectly or did not label the diagram. This once again reflects adequate observation skills but minimal experience on drawing scientific diagrams.

Organizing Information

This area includes skills in making simple interpolations and extrapolations from line graphs, constructing line graphs and performing simple calculations. In comparison to their performance on the other skill areas, the Form 5 students did not do well in this area. In the item that required construction of a line graph, very few students included a title. Even when they
were instructed how to label the axes, about 25% of the students labelled the axes incorrectly. Tamir (1989) reported that the Israeli students did not do well in simple calculations. On the contrary, the Tongan students did relatively well in this area. Two factors could have accounted for this difference. The questions presented to the Tongan students might have been easier than those attempted by the Israeli students, or the Tongan students could have had more experience in these tasks.

Interpreting Information

This area includes skills in classifying organisms and objects, recognizing relationships between variables and trends from graphs and data, identifying inconsistent information and limitations in the design of an experiment and drawing conclusions based on the results of an experiment. In this area, the students performed worst in the skills relating to identifying limitations in the design of an experiment and drawing conclusions based on the results of an experiment. Jacobson and Doran (1988) reported that the American students in Grade Five and Nine performed poorly in explaining, designing experiments and reasoning. Tamir (1989) also reported that the Israeli students did not perform well in inferring and drawing conclusions.
Communication Information

This area includes skills of extracting notes from a text and using contents and index pages to locate information in a text. The students performed relatively very well in this area. Skills of extracting notes and key points from a text book are not explicitly taught as part of the Form 5 General Science curriculum. However, students do learn these skills in other subjects, and also practise them in science while taking notes from texts and while writing laboratory reports.

Difference Between the Performance of Form 3 and Form 5 Students on the Written test

The difference in the performance of Form 3 and Form 5 students on the written test was significant in all the skill areas tested. It could be interpreted that Form 4 and Form 5 curriculum had been successful in producing these differences. However, there are some intriguing aspects of the test results that should be clarified before any such conclusions could be drawn. One is that Form 5 students did best on the areas in which the skills involved are also taught in other subjects. For example, communicating information is taught in English and performing calculations is taught in Mathematics. The other is the generally poor level of skill attainment of the students as revealed by the test results. Their attainment of science process skills after completing Form 4 and Form 5 curriculum has been shown to be very low, particularly in
the areas of planning and problem analysis. Although the growth in performance is significant, the absolute level of performance of the Form 5 students is not satisfactory.

Performance of Form 5 Students on the Practical Test of Science Apparatus Skills

As for the written test, the overall performance of the students on the practical test was very low. However, it had been anticipated that the students would perform better on the practical test due to the lower cognitive demands of the apparatus skills involved in the test.

Measurement of Length, Mass and Temperature

In the task of measuring length, most of the students seemed familiar with the ruler. However, many of them were not precise enough in their measurement. In measuring mass, the students were required to use a spring and a triple-beam balance to measure the mass of two steel bars. On the item that required the students to use a spring balance, the mean score for Form 3 was higher than the mean score for Form 5. Another item on which the mean score for Form 3 was higher than the mean score for Form 5 was the item that required the students to use a measuring cylinder to measure a volume of water. Use of the spring balance and measuring cylinder are listed in the Form 3 General Science curriculum but not in Form 5. The results showed that the
skills learned in Form 3 almost disappeared two years later. This suggests that the students have had insufficient practice to consolidate these skills. For these skills to be acquired by the students, there is a need for considerable practice and feedback (Fitts, 1964).

All the schools included in the study possessed at least one triple-beam balance. However, only about 40% of the students were able to use it to accurately measure the mass of a steel bar.

When measuring temperature, 30% of the students were not sure of the correct end of the thermometer to be immersed in the water. About 20% of the students did not wait to allow the thermometer to settle, and about 40% of them read the scale while the bulb of the thermometer was out of the water. Although all of the schools included in the study possessed thermometers, most of the students reported that this was the first time that they used the instrument.

An interesting characteristic observed during these activities was the students' ignorance of precision when they undertook measurements. Simple precautions that should be taken in order to obtain the most precise readings were ignored by most of the students. For example, more than 70% of the students ignored parallax error when they were reading the scale on the thermometer. More than 70% of the students did not hang the spring balance on a retort stand to stop the motion of the needle so that a more accurate
reading of the mass of the steel bar could be taken. It appears as a lack of proper attitude towards working as scientists. Tongan society has a low level of technology and there are few situations in the world of work where there is a need for people to be precise when dealing with measurements. Children observe this in their society, and when they are involved in science practical activities, they do not consider it critical to be precise when taking measurements.

**Starch Test**

In comparison to their performance on the other tasks, the Form 5 students performed best on the starch test, which includes the skills of performing simple chemical tests. On this task, the students were required to follow a set of instructions to test for the presence of starch in three samples of milk. They were required to test the milk samples with iodine, observe and record observations of colour changes, and draw conclusions based on their observations. The first milk sample contained no starch, the second contained a little starch, and the third sample contained twice the amount of starch put in the second sample. It was expected that the students would detect and report the difference in the intensity of the blue-black colour between the second and the third sample. However, only about 40% of them recorded this difference, and only 50% of this proportion went on to conclude that the third sample had more starch than the second. Discussions with
the students after the test revealed that a considerable number of them did not record the difference in the intensity of the colour because they were not asked to do so. Some thought that it was not important. This reflects the nature of practical work that students perform in their science classes. They are encouraged mainly to look for the expected results of the experiments, to confirm laws and theories taught during science lessons.

Another interesting feature observed on this item was that about 25% of the students recorded their conclusions under the section for their observations. It appeared that many students could not distinguish between observations and conclusions. Most students did relatively well on observation, however, many found it difficult to draw conclusions from their observations. This is consistent with the results of the written test.

**Using a Bunsen Burner to Heat a Test Tube**

On the task which required students to light a Bunsen burner and then use it to heat a test tube containing a liquid, about 70% of the students commenced heating the test tube without trying to change or improve the smokey-yellow flame they obtained. About 90% of them did not wear the safety glasses, 15% did not use the test tube holder and 35% pointed the test tube at some one while heating it. These results indicate little experience in using a Bunsen burner to heat a test tube.
Estimation and Measurement of Volumes of Liquids

In estimating volumes of liquids, the students were required to estimate the volume of water in a jam jar. Many students found this very difficult. They were simply not familiar with liquid volumes. On the other task, the students were instructed to use a measuring cylinder to measure a required amount of water and transfer it to a plastic cup for assessment. Once again, the students were observed to be imprecise while taking measurements.

pH Test

This item also tested the skills in performing simple chemical tests. The students were required to use blue and red litmus papers to identify the contents of four bottles, to determine which bottles contain acid, base or water. The difference between this task and the Starch Test was that this one required the students to outline a plan of their investigation. The students were told that if they did not know what to do, they could open an envelope which contained written instructions for the task. Students who opened the envelope lost marks. About 70% of the students did not open the envelope but only about 15% provided an appropriate plan for the investigation. However, about 60% of the students correctly identified the contents of the bottles. This indicated that many students were able to perform the task but did not know how to plan it. Hackling and Garnett (1990) observed that one of the most
distinctive features of students' problem solving was the limited amount of problem analysis and planning done before manipulating the equipment and collecting data.

**Gender Differences in the Performance of Form 5 Students on the Tests**

When the individual items were grouped into skill areas, it was found that there were no significant differences between the performance of boys and girls in each area.

When comparing the means for the boys and the girls for each of the items of the written test, it was apparent that the boys outperformed the girls on items OE02 and OE07. These items required the students to carry out simple calculations. In the Tongan national examinations, boys have generally performed better than girls in Mathematics. Girls outperformed boys on item OE10. This item required the students to extract important information from a text. This feature has also been observed in the results of the Tongan national examinations over the years, that is the girls have generally performed better than the boys in the language subjects like English and History (Tongan Ministry of Education, 1980; 1981; 1982).

A considerable number of studies have reported the existence of differences in the performances of boys and girls in various areas of science (Johnson & Murphy, 1984;
Kelly, 1978). However, there has been little evidence to support the case for gender differences in the area of science process skills. For example, Jacobson and Doran (1988) reported that the expected male-female differences on the performance of the American students on the science process laboratory skills were almost non-existent. Johnson and Murphy (1984) also observed that there were minimal differences between the performances of boys and girls on science process skills tests.

In the practical test of science apparatus skills, there were no significant differences between the performance of boys and girls in any of the skill areas tested. Physical environment is one of the factors that cause the differences in scientific experiences and interests of boys and girls before they attend high school science (Johnson & Murphy, 1984). The environment of Tonga does not provide the opportunities for the children to be involved with technology-based leisure-time activities like mechanical activities and constructional games for boys, and hand lenses, weighing scales and measuring jugs for girls. The only time for all children to experience and attain these skills is during science practical activities in school. Therefore, although the Tongan boys and girls are involved in different types of leisure-time activities, they do not benefit much from these activities in terms of science investigation skills.
Assessment Procedures in Tonga

In the chief examiner's report on the results of the first Tongan School Certificate Science examination in 1989, he stated that in the questions where the emphasis was on the practical skills, it was apparent that students had performed poorly. The assessment of science practical skills is not emphasized in the prescription for the Tongan School Certificate examination prepared by the Tongan Ministry of Education (1990). It only suggests that some questions may be designed to test the 'appreciation of scientific attitudes and processes and their relevance to situations of social significance' (p. 30). However, two of the seven major objectives of the course are:

(c.) The application of the scientific method: the ability to identify a problem, to bring to bear earlier experience relevant to the problem, to formulate explanations and hypotheses, to test by experiment or other means, to accept, modify, or reject, and to draw conclusions.

(d.) The development of skills appropriate to science: the ability to use scientific equipment accurately, to construct and interpret tables, charts, and graphs, to find relevant information from reference sources (p. 30)

Most of these skills are not assessed in the national examination. Therefore they can only be assessed through practical work performed during the course. The method of assessment of practical work in Tonga is through students' written reports of laboratory work. The Ministry's requirement for practical work is for each student to
complete a certain number of experiments before the student can sit the written examination. An inspector from the Ministry visits the schools to check that the students have completed the required number of experiments. It is obvious from the above discussion that the objectives relating to science investigation skills are not adequately taught or properly assessed in Tonga. This study has shown that attainment of science investigation skills by the Tongan Form 5 students has fallen short of its objectives. Part of this failure is due to the method of assessment employed to evaluate these skills.
CHAPTER 7

Conclusions and Implications

Introduction

This chapter presents the conclusions and implications of the study. First, the findings of the study are summarised in relation to the research questions. This is followed by a discussion of the implications for teaching and for further research.

Conclusions of the Study

The first research question of the thesis sought to identify the science investigation skills and standards of performance that students should attain after completing the Tongan Form 5 General Science program. This question prompted the development of benchmark statements which describe the range of science investigation skills and standards of performance that should be expected of Tongan Form 5 General Science students. These include skills of problem analysis, planning an investigation, manipulating scientific equipment, collecting information, organizing information, interpreting information, and communicating information. The benchmark statements, together with the behavioural objectives for each benchmark are presented in Appendix 1.
The second research question sought to identify the skills that students have attained before and after completing the Tongan Forms 4 and 5 General Science programs. Test instruments were developed and administered to the Tongan Form 3 and Form 5 General Science students.

It was found that in the skill areas tested by the written test of science process skills, most of the students were still working towards the benchmark. Form 5 students performed best in collecting and communicating information and worst in the areas of planning and problem analysis. Very few Form 5 students understood the concepts of hypothesis, independent, dependent and control of variables and how they relate to each other. Most of the laboratory activities that students are doing are of recipe style (Friedler & Tamir, 1984; Tamir & Lunetta, 1981; Tobin, 1986) which provides little opportunity for students to practise planning controlled experiments. Tongan science teachers reported that the Form 5 students received little instructions on types of variables.

About 50% of the Form 5 students performed above benchmark standard in four of the eight objectives assessed by the practical test of science apparatus skills. Most of the Form 5 students were not competent in using a thermometer to measure temperature, a Bunsen burner to heat a test tube and a triple-beam balance to measure the mass of an object. Form 5 students were not precise when taking measurements using a ruler, a spring and a triple-beam
balance, a thermometer and a measuring cylinder. These results suggest that the Tongan Form 5 General Science students do not have adequate practice in working with scientific equipment.

The third research question sought to identify the effectiveness of the Tongan Forms 4 and 5 General Science programs in developing science investigation skills. It was found that Form 5 students performed significantly better than the Form 3 students in these skills, but not up to the standards described in the benchmark statements. It can therefore be argued that the Form 4 and 5 General Science curricula were not effective in developing science investigation skills to the standards outlined in the benchmark statements.

The subsidiary research question of the thesis sought to identify any difference in levels of skill attainment of Tongan Form 5 boys and girls. It was found that there were no significant differences between the performances of boys and girls in any of the skill areas assessed. However, in the analysis of their performance in individual items, boys performed better than girls in the items that required calculations to be performed. Girls performed better than boys in an item that required them to summarize paragraphs from a science text book.
Implications for Teaching

The results of this study indicate that there is a need to improve the level of science investigation skill attainment of the Tongan Form 5 General Science students. The present Tongan Form 4 and 5 General Science programs need to be revised to improve the teaching and learning of science investigation skills.

Most of the available laboratory practicals are of a recipe type (Friedler & Tamir, 1984; Tamir & Lunetta, 1981; Tobin, 1986), in which students have very little opportunity to practise the higher level skills of problem analysis, planning and control of variables. Attempts should be made to produce new laboratory practicals in which students are involved in activities that would enable them to learn the higher science investigation skills.

It has been argued that inquiry oriented laboratory work is cognitively demanding (Friedler & Tamir, 1986; Johnstone & Wham, 1982) and that working memory may be overloaded with all the information needed. Attempts should therefore be made to teach the conceptual knowledge regarding the structure of controlled experiments initially in a non laboratory situation so that the additional burden of working with apparatus is avoided. Apparatus skills such as measurement of length, mass, temperature and volumes of liquids and performing chemical tests could best be developed through practical laboratory activities designed
specifically to develop these skills rather than in practicals designed to develop science concepts.

The last suggestion relates to the assessment of science investigation skills. Laboratory work is only assessed in Tonga through students' written reports of laboratory exercises. In the Tongan School Certificate Science Examination, only a very small proportion of the test is devoted to the assessment of science process skills. Further, existing published tests of science process skills are not available in Tonga. It is suggested here that the method of assessment of practical work in Tonga be reviewed, and the proportion of the questions relating to science process skills in the Tongan School Certificate Science Examination be increased. Form 4 and 5 General Science teachers need to be provided with some of the available tests of science process skills, and be given in-service training on the teaching and assessment of science investigation skills.

**Implications for Further Research**

Failure of the items in the written test of science process skills to fit the Rasch model raises the question of the validity of producing a single test that covers all science process skills. The alternative to this is to produce separate tests for each of the process skill areas. It is recommended from this study that further research to
this end be carried out to determine the appropriateness of producing tests for specific process skill areas rather than a single comprehensive test.

Each of the test instruments developed in this study were trialled twice in high schools in Perth, and once in Tonga. However, there is still room for improvement in the tests. Further research in Tonga should be conducted to evaluate the appropriateness of translating the instruments into Tongan. English and Tongan versions of the tests should be piloted in Tonga to determine which version promotes the best measure of science investigation skills.

The monitoring standards project should be repeated at three to five year intervals to monitor the development of science investigation skills in Tongan Form 5 students as teaching and assessment procedures are modified and improved.


Cambridge, MA: Harvard University Press.


REFERENCES


REFERENCES

46, 5, 22-25.


Oslon, D. R. (1973). What is worth knowing and what can


1. PROBLEM ANALYSIS
The student demonstrates understanding of a problem by developing questions for investigation, formulates hypotheses for testing and predicts the outcome of specific investigations.

Behavioural Objectives
The student is able to:
(a) state the problem to be investigated;
(b) develop questions for investigation when background information is provided;
(c) identify the variables that could influence the phenomenon to be investigated;
(d) identify dependent and independent variables;
(e) develop hypotheses that state the relationship between the dependent and independent variables; and
(f) predict the outcome of particular investigations or events.
2. PLANNING

When investigating a problem the student specifies the experimental conditions (variables) to be controlled, plans procedures for the manipulation and measurement of variables, and selects appropriate instruments and materials for collecting data.

**Behavioural Objectives**

The student is able to:

(a) define variables in operational terms;
(b) identify variables to be held constant;
(c) identify variables to be measured;
(d) explain why particular variables are to be manipulated and others are to be controlled;
(e) describe how and when measurements or collection of particular data should be made and recorded, and select appropriate materials and equipment required for the tasks; and
(g) plan experimental procedures that minimize safety hazards.
3. MANIPULATING EQUIPMENT
The student manipulates scientific equipment to perform laboratory techniques and collect data, and minimizes safety hazards.

Behavioural Objectives
The student is able to:
(a) collect a gas in a container, and transfer a liquid or a solid to a required level in a container;
(b) employ an appropriate filtering technique;
(c) demonstrate safe working techniques while:
   - lighting and adjusting the flame of a Bunsen burner;
   - heating a solid substance;
   - heating a liquid in a test-tube;
   - heating a beaker of water;
(d) use a Celsius laboratory thermometer, a stop-watch or a ticker-timer, a metre stick or flexitape, and a graduated cylinder to make accurate measurements of temperature, time, volume and length.
(e) read a variety of scales and make appropriate settings and adjustments of various scientific equipment.
(f) use an ammeter and a voltmeter to make accurate measurements of current and voltage; and
(g) perform simple chemical tests.
4. COLLECTING INFORMATION

The student observes, describes and measures specific characteristics and changes, and records information in a systematic manner using a variety of formats and metric units.

**Behavioural Objectives**

The student is able to:

(a) correctly follow instructions for the use of unfamiliar techniques or apparatus;

(b) measure in a consistent manner to maximise precision,

(c) identify sources of error and modify procedures to improve accuracy;

(d) record measurements using correct metric units;

(e) construct well-structured tables for recording numerical data;

(f) describe observations of specific characteristics or changes;

(g) record observations using scientific diagrams that include a title, labels and a scale.
5. ORGANIZING INFORMATION

The student transforms information into alternative forms to facilitate its use and interpretation. The student constructs graphs (kite graphs or line graphs) with appropriate scales and labels.

**Behavioural Objectives**

The student is able to:

(a) perform calculations when given the appropriate mathematical expression;

(b) convert different values to the same unit for comparison and computation;

(c) construct kite graphs or line graphs when told which graph type is appropriate for the given data, and which variable is to be plotted on each axis; and

(d) make simple interpolations and extrapolations from line graphs.
6. INTERPRETING INFORMATION

The student interprets trends in and relationships between information in order to classify, hypothesize and generalise.

**Behavioural Objectives**

The student is able to:

(a) suggest explanations for observations;
(b) classify objects, materials and organisms using relevant characteristics;
(c) describe in words, qualitative or quantitative relationships between variables, and trends from graphs and tabulated data;
(d) recognize inconsistent and anomalous information and propose a reasonable explanation for this information;
(e) recognize limitations in the design of the experiment that influence the reliability of the data gathered;
(f) propose a conclusion or generalisation based on results of an investigation;
(g) formulate new questions or problems from the outcomes of an experiment; and
(h) explain whether an hypothesis, inference or generalisation is supported by evidence from an investigation.
7. COMMUNICATING INFORMATION

The student extracts from a scientific report, article, text or reference book, information directly relevant to an aspect of student inquiry. The student prepares reports of practical investigations such that results and conclusions are logically presented.

Behavioural Objectives

The student is able to:

(a) use contents and index pages correctly to locate information from texts and reference books;

(b) extract, as notes, key points of information from a newspaper article, magazine article, information sheet, text, reference book or audio visual;

(c) prepare reports of library research that are concise, well-structured with contents page and subheadings, and present information in an appropriate sequence;

(d) correctly use specific scientific terminology; and

(e) prepare a well-structured report of a laboratory investigation that states the aim, procedure, results and conclusions.
APPENDIX 2

Objectives Tested in the Written Test of Science Process Skills

<table>
<thead>
<tr>
<th>Benchmarks and Objectives</th>
<th>Question number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M.C.Q.</td>
</tr>
<tr>
<td><strong>1. Problem Analysis</strong></td>
<td></td>
</tr>
<tr>
<td>Identify the variables</td>
<td>3</td>
</tr>
<tr>
<td>that could influence the</td>
<td></td>
</tr>
<tr>
<td>phenomenon to be</td>
<td></td>
</tr>
<tr>
<td>investigated.</td>
<td></td>
</tr>
<tr>
<td>Identify dependent and</td>
<td>4</td>
</tr>
<tr>
<td>independent variables.</td>
<td></td>
</tr>
<tr>
<td>Identify or develop</td>
<td>1</td>
</tr>
<tr>
<td>hypothesis that state the</td>
<td></td>
</tr>
<tr>
<td>relationship between</td>
<td></td>
</tr>
<tr>
<td>dependent and independent</td>
<td></td>
</tr>
<tr>
<td>variables.</td>
<td>5</td>
</tr>
<tr>
<td><strong>2. Planning</strong></td>
<td></td>
</tr>
<tr>
<td>Identify variables</td>
<td>2</td>
</tr>
<tr>
<td>to be held constant.</td>
<td>6</td>
</tr>
<tr>
<td>Explain why particular</td>
<td></td>
</tr>
<tr>
<td>variables are to be</td>
<td></td>
</tr>
<tr>
<td>manipulated and others</td>
<td></td>
</tr>
<tr>
<td>to be controlled.</td>
<td></td>
</tr>
<tr>
<td>Describe how and when</td>
<td>7</td>
</tr>
<tr>
<td>measurements should be</td>
<td></td>
</tr>
<tr>
<td>made and recorded.</td>
<td></td>
</tr>
</tbody>
</table>
### 4. Collecting Information

<table>
<thead>
<tr>
<th>Construct well structured tables for recording data.</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Describe observations of specific characteristics or changes.</td>
<td>8</td>
</tr>
<tr>
<td>Record observations using scientific diagrams that include a title, labels and a scale.</td>
<td>9</td>
</tr>
</tbody>
</table>

### 5. Organizing Information

<table>
<thead>
<tr>
<th>Perform calculations when given the appropriate mathematical expression.</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convert different values to the same unit for comparison and computation.</td>
<td>2</td>
</tr>
<tr>
<td>Construct appropriate graphs when told which graph type is to be used and which variable is plotted on each axis</td>
<td>3</td>
</tr>
<tr>
<td>Make simple interpolations and extrapolations from line graphs.</td>
<td>8</td>
</tr>
</tbody>
</table>
### 6. Interpreting Information

<table>
<thead>
<tr>
<th>Activity</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classify or identify objects, materials and organisms on basis of specific descriptions or data.</td>
<td>10</td>
</tr>
<tr>
<td>Describe qualitative and quantitative relationship between variables, and trends from graphs and tabulated data.</td>
<td>9</td>
</tr>
<tr>
<td>Recognize inconsistent and anomalous information and propose a reasonable explanation for this information.</td>
<td>4</td>
</tr>
<tr>
<td>Recognize limitations in the design of an experiment that influence the reliability of the data gathered.</td>
<td>13</td>
</tr>
<tr>
<td>Propose a conclusion or generalization based on results of an investigation.</td>
<td>14</td>
</tr>
</tbody>
</table>

### 7. Communicating Information

<table>
<thead>
<tr>
<th>Activity</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use contents and index pages correctly to locate information from texts and reference books.</td>
<td>1</td>
</tr>
<tr>
<td>Extract, as notes, key points of information from a newspaper article, magazine article, information sheet, text, reference books or audio visual.</td>
<td>10</td>
</tr>
</tbody>
</table>
APPENDIX 3

The Written Test of Science Process Skills

TEST OF SCIENCE PROCESS SKILLS

This is not a test of the factual knowledge from your last science topic. It is a test of your skills of working with scientific data.

INSTRUCTIONS

1. Answer the multiple choice questions on the separate sheet provided.
2. Answer the other questions in the spaces provided in the booklet.

NAME :

Date of birth __________________________

Form / Class __________________________

School _______________________________
Questions 1 and 2 refer to the following situation.

Mary wanted to identify the factors that affect the growth of bread mould. She grew the mould in the dark, in nine containers containing the same amount and type of nutrients. Three containers were kept at 0 °C, three were kept at room temperature (about 27 °C), and three were kept at 90 °C. The containers were examined and the growth of the bread mould was recorded at the end of four days.

1. Which of the following is her hypothesis?
   a. The growth of bread mould is affected by the type of nutrient used.
   b. The growth of bread mould is affected by the temperature.
   c. The growth of bread mould is affected by the amount of nutrient used.
   d. The growth of bread mould is affected by the amount of light.

2. Which of the following variables is controlled (held constant) in this bread mould experiment?
   a. Temperature of the containers
   b. Amount of nutrients
   c. Temperature of the bread mould
   d. Growth of bread mould

3. Jane wanted to find out what might affect the growth of bean plants. She had ten bean plants and all of them were kept moist throughout the experiment. She put five of the plants in a warm sunny window. She put the other five plants in a dark refrigerator. She measured the height of all plants after one week.

Which of the following variables might cause differences in growth of the plants in this experiment?
   a. Temperature only
   b. Moisture and light
   c. Time only
   d. Temperature and light
Questions 4-6 refer to the following situation.

A forest fire destroyed trees in a large area. The forest rangers used the burned area to study the effect of different types of grasses on soil erosion. They chose ten plots of ground that are the same size. Each of these plots receive the same amount of sun, have the same kind of soil and the same slope. The rangers planted each plot with a different type of grass. Measurements of soil erosion were made every week for the entire summer.

4. What is the independent variable (the variable manipulated by the scientist) in this experiment?
   a. The size of the plots
   b. The types of grasses
   c. The amount of soil erosion
   d. The type of soil in the plots

5. What hypothesis is being tested in this study?
   a. The amount of soil erosion depends on the type of grass planted.
   b. Slope of the land affects soil erosion.
   c. Burned areas have greater erosion than forested areas.
   d. Planting grass will reduce the amount of soil erosion.

6. Which of the following variables is NOT controlled (held constant) in this study?
   a. The size of the plots
   b. The type of soil in the plots
   c. The amount of soil erosion
   d. The amount of sun the plots receive
7. Suppose you wanted to test the following hypothesis:

*The hotter the water, the faster sugar will dissolve.*

Which would be the best plan to test the hypothesis?

a. Set up two beakers of water; one at a temperature of 20 °C, and the other at 30 °C. Put one teaspoon of sugar in each beaker and use Benedict's solution to test for the presence of sugar.

b. Set up four beakers of water; one at 20 °C, one at 40 °C, one at 60 °C, and one at 80 °C. Put one teaspoon of sugar in each beaker and measure the time taken for the sugar to dissolve.

c. Set up three beakers of water. Put one teaspoon of sugar in each beaker. Heat beaker one over low heat, beaker two over moderate heat and beaker three at high heat. Measure the time taken for the sugar to dissolve.

d. Set up four beakers of water. Put one teaspoon of sugar in beaker one, two teaspoons in beaker two, three in beaker three and four in beaker four. Measure the time taken for the sugar to dissolve.
8. A thumb tack was used to punch a small hole in the bottom of a paper cup. The hole was covered up and the cup was filled with water. When a signal was given, the hole was uncovered and the time for water to empty from the cup was measured. Every 20 seconds the height of the water left in the cup was recorded. The height of water in the cup was plotted against time.

If the height of the water in the cup was 10 centimetres to start with, use the graph to predict how long it would take to empty. (NB. Do not draw on this graph)

a. 0 seconds  
b. 95 seconds  
c. 100 seconds  
d. 120 seconds
9. The following graph shows the relationship between plant growth and light intensity.

From the graph, the relationship between light intensity and plant growth can be stated:

a. As light intensity increases, plant growth increases.

b. As plant growth increases, light intensity increases to a point, then decreases.

c. As light intensity increases, plant growth increases to a point, then decreases.

d. As plant growth increases, light intensity increases.
10. The following is a key used to classify animals on planet Phool.

1. Has 2 antennae....................go to 2
   Has more than 2 antennae........go to 3

2. Has 4 legs.......................go to 4
   Has less than 4 legs ..........go to 5

3. Has no legs.....................go to 6
   Has legs........................go to 7

4. Has pointed ears............it is a TUPP
   Has rounded ears............it is a GSER

5. Has one eye.................it is a TFUM
   Has more than one eye ....it is a BUHKN

6. Spotted body ..............it is a KRPT
   Not spotted ...............it is a KLRP

7. Has hair ................it is a YOUKL
   Has no hair..............it is a BUKL

Below is a diagram of an animal found on the planet

The animal is a ...

A. BUHKN
B. KRPT
C. GSER
D. YOUKL
11. A student tested the boiling point, freezing point and density of a pure sample of an unknown substance. The student's results are as follows:

<table>
<thead>
<tr>
<th>TEST</th>
<th>RESULTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boiling point</td>
<td>81 °C</td>
</tr>
<tr>
<td>Freezing point</td>
<td>5.6 °C</td>
</tr>
<tr>
<td>Density</td>
<td>0.88 g/cm³</td>
</tr>
</tbody>
</table>

Use the table below to identify the unknown substance.

Is it chemical A, B, C or D?

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Density (g/cm³)</th>
<th>Freezing Point (°C)</th>
<th>Boiling Point (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.88</td>
<td>5.5</td>
<td>80.0</td>
</tr>
<tr>
<td>B</td>
<td>1.50</td>
<td>-63.5</td>
<td>61.2</td>
</tr>
<tr>
<td>C</td>
<td>0.78</td>
<td>6.5</td>
<td>80.7</td>
</tr>
<tr>
<td>D</td>
<td>0.89</td>
<td>16.3</td>
<td>286.0</td>
</tr>
</tbody>
</table>
TEST OF SCIENCE PROCESS SKILLS

PART B

This part of the test contains 14 open-ended questions

Write your answers in this booklet, in the spaces provided

Name

Date of birth

Form / Class

School
1. Below is the contents page from a science book.

<table>
<thead>
<tr>
<th>CONTENTS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 1</td>
<td></td>
</tr>
<tr>
<td>1 How Your Body Works</td>
<td>1</td>
</tr>
<tr>
<td>2 Body Transport Systems</td>
<td>24</td>
</tr>
<tr>
<td>3 Body Control</td>
<td>43</td>
</tr>
<tr>
<td>4 Disease</td>
<td>65</td>
</tr>
<tr>
<td>Chapter 2</td>
<td></td>
</tr>
<tr>
<td>5 Materials from the Earth</td>
<td>87</td>
</tr>
<tr>
<td>6 Making New Materials</td>
<td>112</td>
</tr>
<tr>
<td>7 Acids and Bases</td>
<td>139</td>
</tr>
<tr>
<td>8 Managing our Resources</td>
<td>163</td>
</tr>
<tr>
<td>Chapter 3</td>
<td></td>
</tr>
<tr>
<td>9 The Record in the Rocks</td>
<td>188</td>
</tr>
<tr>
<td>Chapter 4</td>
<td></td>
</tr>
<tr>
<td>10 Inheritance - Searching for Patterns</td>
<td>211</td>
</tr>
<tr>
<td>Chapter 5</td>
<td></td>
</tr>
<tr>
<td>11 Motion</td>
<td>236</td>
</tr>
<tr>
<td>12 Using Energy</td>
<td>258</td>
</tr>
<tr>
<td>13 Energy Alternatives</td>
<td>288</td>
</tr>
<tr>
<td>Answers to Check-Ups</td>
<td>319</td>
</tr>
<tr>
<td>Science Words</td>
<td>330</td>
</tr>
<tr>
<td>Index</td>
<td>337</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>342</td>
</tr>
</tbody>
</table>

Where in the book would you look for the following information?

(a) Energy in an electrical circuit  

(b) Indicators for bases  

(c) Bacteria
2. The relationship between velocity, distance travelled, and time, for an object travelling in a straight line at a constant speed is given by the following expression.

\[
\text{Velocity} = \frac{\text{distance}}{\text{time}}
\]

\[v = \frac{s}{t}\]

How many minutes will it take for a car which is travelling in a straight line at a constant speed of 20 km/hr to cover 8 Km?

From the equation above, \[t = \frac{s}{v}\]
3. An experiment was planned to find out how the height reached by a rocket changed as the launching speed was increased. The results of the experiment are recorded below.

<table>
<thead>
<tr>
<th>Launching Speed</th>
<th>Maximum Height</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 metre per second</td>
<td>25 metres</td>
</tr>
<tr>
<td>2 metres per second</td>
<td>50 metres</td>
</tr>
<tr>
<td>3 metres per second</td>
<td>100 metres</td>
</tr>
<tr>
<td>4 metres per second</td>
<td>200 metres</td>
</tr>
<tr>
<td>6 metres per second</td>
<td>500 metres</td>
</tr>
</tbody>
</table>

In the grid provided below, draw a line graph to represent the data above. Plot launching speed on the horizontal axis.
The effect of exercise on pulse rate was studied by a science class. Students did different numbers of jumps and then measured pulse rate. Group one jumped five times, group two jumped ten times, group three jumped fifteen times, and so on. Their results are recorded in the table below.

<table>
<thead>
<tr>
<th>Group</th>
<th>Number of jumps</th>
<th>Pulse rate (pulses/minute)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>60</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>64</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>70</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>66</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>80</td>
</tr>
</tbody>
</table>

Which of the above group results seem to be out of step with the others?

Give a reasonable suggestion why this group might have recorded this unusual result.
5. Coffee contains caffeine, which is a stimulant drug. Other stimulant drugs prescribed by doctors often cause an increase in pulse rate.

Propose an hypothesis regarding the possible relationship between drinking coffee and pulse rate.

6. Paul wanted to determine the volume of three objects. He took readings of the mass of the objects in air and in water. He then subtracted the values in water from the values in air to determine the volume of the objects. The results are recorded below.

First object: mass in water = 20 g, mass in air = 40 g,
Second object: mass in air = 42 g, mass in water = 30 g,
Third object: mass in water = 36 g, mass in air = 41 g,

Construct a suitable table for the data above, and record the values for the masses and volumes (cm$^3$) in it.
7. The density of an object can be calculated using the expression:

\[
density = \frac{mass}{volume}
\]

Calculate the mass of an object with density $2 \text{ g/cm}^3$ and volume of $10 \text{ cm}^3$ (show your work).

Questions 8–9 refer to the following:

On your desk there are two different fruits. Observe them very carefully.

8. List five ways in which the cut surfaces are different

9. Draw a scientific diagram of the cut surface of the orange.
10. Below are some paragraphs taken from a science book.

Pick out the main ideas and write them as a set of notes.

"Acids which usually react quickly with substances are called strong acids. For example, hydrochloric acid, sulfuric acid, and nitric acid are all strong acids. Acids which react slowly are called weak acids. Most of the acids found in living (or recently living) things are weak acids. For example, acetic acid (from wine), citric acid (from citrus fruits), and lactic acid (in yoghurt).

There are also strong and weak bases. For instance, sodium hydroxide is a strong base, but ammonia is a weak base.

Be careful not to confuse the terms strong and weak with the terms concentrated and dilute. Strong and weak refer to the type of acid. Concentrated and dilute refer to the amounts of water that have been added to the acids. For example, it is possible to have a concentrated strong acid and a dilute strong acid. The difference is due to the water that has been added to dilute the acid. Also, there are concentrated weak acids and dilute weak acids. Later in this chapter, you will learn more about strong and weak acids."

Notes

_________________________________________________________________

_________________________________________________________________

_________________________________________________________________

_________________________________________________________________

_________________________________________________________________

_________________________________________________________________

_________________________________________________________________

_________________________________________________________________
Questions 11-12 refer to the following situation

An investigation was planned to find which laundry detergent was best for removing grass stains from clothing. Each detergent was mixed with warm water and used to scrub a piece of grass-stained material. The material was scrubbed for one minute. After scrubbing, the amount of stain remaining on the material was noted.

11. Which is the dependent variable (the variable that responds to changes in the independent variable) in this investigation?


12. Identify one additional variable that should have been controlled in this investigation. Explain why you think it is important to control this variable?
13. A scientist performed the following experiment to test an hypothesis:

The hypothesis was:

*Left-handed people play better tennis than right-handed people.*

The experiment:

A list of all the tennis players in the country was compiled. The scientist then randomly selected from the list one left-handed and one right-handed player. A match between the two players was won by the left-handed player.

Question:

Write down all the faults in the experimental design that you can think of.
14. A student wanted to find out whether crabs prefer to stay in the light or in the dark. He set up two trays as shown below.

He then placed 5 crabs in the light and five crabs in the dark. Every minute, he counted the number of crabs in the light and the number of crabs in the dark. The results are recorded in the table below.

<table>
<thead>
<tr>
<th>Time (minutes)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of crabs in the light</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>No of crabs in the dark</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>

Write a conclusion to this investigation.

How does the results support your conclusion?
APPENDICES

APPENDIX 4

Marking Key for the Written Test of Science Process Skills

TEST OF SCIENCE PROCESS SKILLS

MARKING KEY

SECTION A.

<table>
<thead>
<tr>
<th>QUESTIONS</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANSWERS</td>
<td>B</td>
<td>B</td>
<td>D</td>
<td>B</td>
<td>A</td>
<td>C</td>
<td>B</td>
<td>C</td>
<td>C</td>
<td>D</td>
<td>A</td>
</tr>
</tbody>
</table>

SECTION B.

<table>
<thead>
<tr>
<th>QUESTIONS</th>
<th>MARKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question 1.</td>
<td>Total = 3</td>
</tr>
<tr>
<td>(a) Chapter 5</td>
<td></td>
</tr>
<tr>
<td>or Page 258 or 288</td>
<td></td>
</tr>
<tr>
<td>or Topic 12 or 13</td>
<td>1</td>
</tr>
<tr>
<td>(b) Chapter 2</td>
<td></td>
</tr>
<tr>
<td>or Page 139</td>
<td></td>
</tr>
<tr>
<td>or Topic 7</td>
<td>1</td>
</tr>
<tr>
<td>(c) Chapter 1</td>
<td></td>
</tr>
<tr>
<td>or Page 65</td>
<td></td>
</tr>
<tr>
<td>or Topic 4</td>
<td>1</td>
</tr>
</tbody>
</table>

Question 2.                 | Total = 2 |
\[ t = \frac{s}{v} \] 1/2 
\[ t = \frac{8}{20} \] 1/2 
\[ t = 0.4 \text{ hr} \] 1/2 
\[ t = 24 \text{ min.} \] 1
Question 3.  

Graph of Launching Speed Vs Maximum Height

(a) Title
(b) Labelling of axes
   - correct variable in each axis
   - correct units
(c) Scale intervals
   - each axis is scaled so the graph takes a large area of the grid
   - scale intervals are of equal value
(d) Marking of points
   - 5 points correct
   - 3 or 4 points correct
   - less than 3
(e) Line connecting points
   - smooth curve
   - very rough curve

Total = 7
Question 4.  
Group with unusual results - Group 4  
Explanation:  
- a relatively fit group  
- error in counting of..  
  number of jumps  
  number of pulses per minute  

Question 5.  
Hypothesis:  
Drinking coffee causes an increase in pulse rate  
(or different wording with the same meaning)  

Question 6.  
Masses and Calculated Volumes of Objects

<table>
<thead>
<tr>
<th>Object</th>
<th>Mass in air (g)</th>
<th>Mass in water (g)</th>
<th>Volume (cm$^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Object</td>
<td>40</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Second Object</td>
<td>42</td>
<td>30</td>
<td>12</td>
</tr>
<tr>
<td>Third Object</td>
<td>41</td>
<td>36</td>
<td>5</td>
</tr>
</tbody>
</table>

Title  
Correct labelling of..  
  columns  
  rows  
Correct units  
  - placed next to the column headings  
  - placed together with values  
Correct values  
  - all rows correct  
  - 2 rows correct  
  - correct rows but no column for volume  
  - 1 row correct
Question 7.

mass = density \times volume

mass = 2 \times 10

mass = 20 g

Total = 2

1/2

1/2

1

Question 8.

Five ways in which the fruits pieces are different

(a) Size
    mentions which one is bigger

(b) Number of segments
    gives correct number of segments in each fruit

(c) Colour
    also gives colour of each

(d) Number of seeds
    also gives the number observed in each

(e) Texture
    also describes which one is smooth and
    which one is rough

(Other reasonable observed differences are accepted but only five
should be marked)

Total = 5

1/2

1/2

1/2

1/2

1/2

1/2

Question 9.

Title

Labels

two or more labels correct

one correct

Scale

correct

wrong

no scale mentioned

Accuracy

tidy and correct number of segments (10)

very rough but right number of segments

tidy but wrong number of segments

Total = 4

1

1

1/2

1

1/2

0

1

1/2

1/2
Question 10.  

Notes:
- Strong acids react quickly with substances  
  e.g. hydrochloric acid  
  - 1
- Weak acids react slowly with substances  
  e.g. citric acid  
  - 1
- There are strong and weak bases  
  - 1
- Strong and weak refer to the type of acid  
  while concentrated and dilute refer to the  
  amounts of water added to the acids.  
  - 1
- Any additional point  
  - 1/2

Question 11.  

Dependent variable:
Amount of stain remaining on the material
after scrubbing (or cleanliness of material)  
  - 2

Question 12.  

Additional variable:
* Amount of detergent (or water) - the concentration of  
  the detergent should be kept constant because even for  
  the same detergent, a more concentrated solution will  
  remove dirt better than a dilute solution.

* Temperature - Generally, the rate of reactions are  
  higher in high temperatures than in low temperatures.  
  It will affect the rate at which each detergent  
  removes stain from clothing.

* Mode of scrubbing - differences in method and weight of  
  scrubbing will affect the amount of stain removed by  
  each detergent.

* Rate of scrubbing - the number of times the stain has  
  been scrubbed in one minute also affects the amount  
  of stain left.

One mark should be given for any of the variables above  
and two marks for an adequate explanation.
Question 13.  

Faults in the experiment:

- sample size too small  
  2

- failed to control interfering variables such as:
  - sex  
  1
  - experience  
  1
  - or any other two variables not controlled


Question 14.  

Conclusion:

Crabs prefer to stay in the dark rather than in the light  
  2

Results

After 10 minutes, there were more crabs (8) in the dark than in the light (2).  
  2
APPENDIX 5

Objectives Tested in the Practical Test

<table>
<thead>
<tr>
<th>OBJECTIVES</th>
<th>STATION NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurement of length and mass</td>
<td>1</td>
</tr>
<tr>
<td>Performing simple chemical tests</td>
<td>2 &amp; 5</td>
</tr>
<tr>
<td>Using a Bunsen burner to heat a liquid</td>
<td>3A</td>
</tr>
<tr>
<td>Using a laboratory thermometer</td>
<td>3B</td>
</tr>
<tr>
<td>Estimation of volumes of liquids</td>
<td>4A</td>
</tr>
<tr>
<td>Using measuring cylinders to measure volumes of liquids</td>
<td>4B</td>
</tr>
</tbody>
</table>
APPENDICES

APPENDIX 6

The Practical Test of Science Apparatus Skills

TEST OF SCIENCE APPARATUS SKILLS.

Name: ____________________________________________________________

Birth Date: ____ / ____ / ______

Form / Class: ______________________________

School: ______________________________

This is a practical science test. It is a test of your skills in using scientific equipment to carry out experiments.

INSTRUCTIONS:

1. You will be asked to do some experiments at five stations round the room. The materials you need have been set out for you at these stations.

2. This paper explains what you should do at each station. You are required to follow the instructions and answer the questions in the spaces provided.

3. You will be given 7 minutes to complete the experiment(s) at each of the stations.

4. After completing the experiment(s) at each station, the supervisor will ask you to clean up the station and return the equipment to its original condition. Please do this before you move on to the next station.

5. When the supervisor asks you to change stations you should move in a clockwise direction.
STATION 1.

Measurement of Length and Mass.

(a) Measure as accurately as possible the length of the two sticks provided.

(i) Stick A.

Use the 30 cm rule to accurately measure the length between the red marks on stick A.

length = _____________________

(ii) Stick B.

Use the 1 M rule to accurately measure the length between the red marks on stick B.

length = _____________________

(b) Measure the mass of the steel bars on the table.

(i) Steel bar A.

Use the 200 g spring balance to accurately measure the mass of steel bar A.

mass = _____________________

(ii) Steel bar B.

Use the triple-beam balance to accurately measure the mass of steel bar B.

mass = _____________________
STATION 2.

Starch Test.

It is rumoured that starch has been added to a brand of milk to make it appear thicker. You are provided with three samples of milk A, B and C. At least one of the samples of milk contains starch.

Iodine solution is used to test for the presence of starch. Iodine will turn blue-black in the presence of starch.

(a) Shake the milk containers then half fill the cups labelled A, B, C with milk from the containers.

(b) Test the milk in the cups using the iodine solution and record your observations and conclusions below.

(c) Record your observations.

Sample A. ____________________________________________________________
Sample B. ____________________________________________________________
Sample C. ____________________________________________________________

(d) What are your conclusions?

Sample A. ____________________________________________________________
Sample B. ____________________________________________________________
Sample C. ____________________________________________________________

Please empty the cups that you used and then put them into the tray labelled "Used Cups".
 Part I. Using a Bunsen burner to heat a liquid.

This exercise requires you to light a Bunsen burner and heat a liquid in a test tube over the flame.

(a) Use the matches provided to light the Bunsen burner

(b) Pour some of the liquid from Beaker A into the test tube and heat it over the flame until the colour of the liquid turns orange.

 Part II. Using the Thermometer.

(c) Use the thermometer to accurately measure the temperature of the water in the beaker.

\[\text{Temperature} = \phantom{000}\]
Measurement and estimation of volumes of liquids.

Part A.

(i) Estimate the volume of water in the test tube.

Volume is about _________.

(ii) Estimate the volume of water in the jam jar.

Volume is about _________.

Part B.

Use the equipment on the table to carry out the following measurements.

Use the felt-tipped pen provided to write your name on the labels on the plastic cups.

(i) Use the 250 mL cylinder to accurately measure 165 mL of water from the container "X" and pour it into plastic cup 1.

(ii) Use the 100 mL cylinder to accurately measure 27 mL of water from container "X" and pour it into plastic cup 2.

Place the plastic cups in the tray for assessment.
STATION 5.

pH Test.

Litmus paper is an indicator used in the identification of acids and bases. Blue litmus paper turns red when dipped in an acid. Red litmus paper turns blue when dipped in a base.

You are provided with four dropper bottles, labelled A, B, C, and D. Your task is to use the litmus paper to determine which bottles contain acid, base or water. Water is neither an acid nor a base.

(a) How would you find this out? Outline your plan.

If you are not sure how to do this task open the envelope marked "Plan". Please do not open the envelope if you know what to do.

Carry out your experiment.

(b) Record your observations.

Bottle A.

Bottle B.

Bottle C.

Bottle D.

(c) What are your conclusions?

Bottle A.

Bottle B.

Bottle C.

Bottle D.

Put the cups that you used into the tray labelled "Used Cups"
Observation Sheet for Station 3

Name of Student ____________________________

<table>
<thead>
<tr>
<th>What to observe</th>
<th>Marking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td><strong>Part I</strong></td>
<td></td>
</tr>
<tr>
<td>lighting burner</td>
<td>not adjust air hole</td>
</tr>
<tr>
<td></td>
<td>not wear safety glasses</td>
</tr>
<tr>
<td>amount of solution in test tube</td>
<td>&gt; 1/2</td>
</tr>
<tr>
<td>heating test tube (tt)</td>
<td>a yellow flame</td>
</tr>
<tr>
<td></td>
<td>not use tt holder</td>
</tr>
<tr>
<td></td>
<td>tt pointed at someone</td>
</tr>
<tr>
<td></td>
<td>tt not at an angle</td>
</tr>
<tr>
<td></td>
<td>not prevent bumping</td>
</tr>
<tr>
<td><strong>Part II</strong></td>
<td></td>
</tr>
<tr>
<td>time allowed for thermometer to settle</td>
<td>&lt; 15 sec</td>
</tr>
<tr>
<td>position of bulb while equilibrating</td>
<td>touch beaker</td>
</tr>
<tr>
<td>position of bulb when reading scale</td>
<td>out of water</td>
</tr>
<tr>
<td>reading of scale</td>
<td>ignore parallax error</td>
</tr>
<tr>
<td>taking care of thermometer</td>
<td>left in beaker</td>
</tr>
</tbody>
</table>

Actual temperature of water = _______
Plan required in Station 5

PLAN

(a) Put some few drops (about 2 mL) from the dropper bottle to its corresponding plastic cup.

(b) Dip a piece of red litmus paper and a piece of blue litmus paper into the solution in each cup and record any change in the colour of the litmus paper.
## APPENDIX 7

### Marking Key for the Practical Test of Science Apparatus Skills

#### MARKING KEY FOR THE TEST OF SCIENCE APPARATUS SKILLS

#### Station 1.

<table>
<thead>
<tr>
<th>Item No</th>
<th>Answer</th>
<th>Marking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>$&gt; 21.35$</td>
<td></td>
</tr>
<tr>
<td>1(a) (ii)</td>
<td>53.2 cm</td>
<td>$&lt; 53.0$</td>
</tr>
<tr>
<td></td>
<td>$&gt; 53.4$</td>
<td></td>
</tr>
<tr>
<td>1(b) (i)</td>
<td>34.5g</td>
<td>$&lt; 23.7$</td>
</tr>
<tr>
<td></td>
<td>$&gt; 25.3$</td>
<td></td>
</tr>
<tr>
<td>1(b) (ii)</td>
<td>484.45g</td>
<td>$&lt; 484.15$</td>
</tr>
<tr>
<td></td>
<td>$&gt; 484.75$</td>
<td></td>
</tr>
</tbody>
</table>

Units: 1 or 2 incorrect units or symbols used ................. lose 1 mark

more than 2 incorrect units or symbols used ........ lose 2 marks

#### Station 2.

<table>
<thead>
<tr>
<th>Item No</th>
<th>Answer</th>
<th>Marking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>(c)</td>
<td>sample A: no change in colour or brown/yellow colour</td>
<td>wrong</td>
</tr>
<tr>
<td></td>
<td>sample B: blue-black colour</td>
<td>wrong</td>
</tr>
<tr>
<td></td>
<td>Sample C: blue-black colour and darker than B</td>
<td>wrong</td>
</tr>
<tr>
<td></td>
<td>Sample C: blue-black colour and darker than B</td>
<td>wrong</td>
</tr>
<tr>
<td>(d)</td>
<td>sample A: no starch</td>
<td>wrong</td>
</tr>
<tr>
<td></td>
<td>sample b: contains starch</td>
<td>wrong</td>
</tr>
<tr>
<td></td>
<td>sample C: contains starch and has more starch than B</td>
<td>wrong</td>
</tr>
</tbody>
</table>
Station 3.

Part I

(a) Lighting the burner:
   Adjusted air hole ..................................... 1
   Not adjusted air hole .................................... 0
   Wear safety classes ..................................... 1

(b) Heating liquid in the test tube:
   Blue flame
   Amount of liquid < 1/2
   Used test tube holder
   Test tube not pointed at anyone
   Test tube at an angle
   Prevented pumping
   - performed all ......................................... 3
   - performed 3 or 4 of them ............................ 2
   - performed 1 or 2 of them ............................ 1

Part II

(c) Procedure of using thermometer:
   Knew the correct end (with bulb) to use
   Allowed > 15 minutes for thermometer to settle
   Thermometer's bulb did not touch beaker
   Bulb submerged while reading the scale
   Avoided parallax error while reading the scale
   Left the thermometer in a safe place
   - performed all ......................................... 3
   - performed 3 or 4 of them ............................ 2
   - performed 1 or 2 of them ............................ 1

Record and reading of temperature:

<table>
<thead>
<tr>
<th>What to observe</th>
<th>Marking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>accuracy of temperature reported</td>
<td>error is &gt; 1 °C</td>
</tr>
<tr>
<td>report of temperature</td>
<td>wrong or no units</td>
</tr>
</tbody>
</table>
### Station 4

<table>
<thead>
<tr>
<th>Item NO</th>
<th>Answer</th>
<th>Marking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(i)</td>
<td>15 mL</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt; 11</td>
<td>11 * - 18</td>
</tr>
<tr>
<td></td>
<td>&gt; 18</td>
<td></td>
</tr>
<tr>
<td>(ii)</td>
<td>125 mL</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt; 94</td>
<td>94 * - 156</td>
</tr>
<tr>
<td></td>
<td>&gt; 156</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(i)</td>
<td>165 mL</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt; 159.00</td>
<td>159.00 * - 171.00</td>
</tr>
<tr>
<td></td>
<td>&gt; 171.00</td>
<td></td>
</tr>
<tr>
<td>(ii)</td>
<td>27 mL</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt; 24.00</td>
<td>24.00 * - 30.00</td>
</tr>
<tr>
<td></td>
<td>&gt; 30.00</td>
<td></td>
</tr>
</tbody>
</table>

Units for Part A:
One or both units incorrect - lose one mark
Station 5.

(a) Plan:
Mentioned that there is a need to test each solution with both types of litmus paper ....... 1

(b) Observations:

A - red litmus no change
   blue litmus to red
B - red litmus to blue
   blue litmus no change
C - red litmus no change
   blue litmus no change
D - red litmus no change
   blue litmus to blue

- all 8 correct .................................. 4
- 6 or 7 correct .................................. 3
- 4 or 5 correct ................................. 2
- 2 or 3 correct ................................. 1
- 0 or 1 correct ................................. 0

(c) Conclusion:
A - an acid ........................................ 1
B - a base ........................................ 1
C - is water ...................................... 1
D - an acid ........................................ 1
## APPENDIX 8

### Equipment and Materials List for the Practical Test

**EQUIPMENT AND MATERIALS LIST FOR THE TEST OF SCIENCE APPARATUS SKILLS.**

<table>
<thead>
<tr>
<th>Station</th>
<th>Quantity</th>
<th>Item</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1 mm division 30 cm rule</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1 cm division metre rule</td>
<td>with only one side graduated</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>40 cm stick</td>
<td>with two red marks 25.45 cm apart</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>80 cm stick</td>
<td>with two red marks 48.2 cm apart</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0-200 g spring balance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>triple-beam balance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>25 g steel bar (approx)</td>
<td>Labelled &quot;A&quot;</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>443 g steel bar (approx)</td>
<td>Labelled &quot;B&quot;</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>retort stand</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>18</td>
<td>Plastic cups (small)</td>
<td>Prepare a set of 3 cups labelled A, B and C for each student and one set as a spare</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>dropper bottle of Iodine solution</td>
<td>Labelled: &quot;Iodine&quot; and &quot;Poison&quot;</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1 L plastic container</td>
<td>Labelled:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&quot;Milk A&quot; (contains 200 mL milk)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>&quot;Milk B&quot; (contains 200 mL milk and 2 tsp. starch)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>&quot;Milk C&quot; (contains 200 mL milk and 6 tsp. starch)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Sponge</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Tray</td>
<td>Labelled: &quot;Used Cups&quot;</td>
</tr>
<tr>
<td>Station</td>
<td>Quantity</td>
<td>Item</td>
<td>Comments</td>
</tr>
<tr>
<td>---------</td>
<td>----------</td>
<td>------</td>
<td>----------</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>Bunsen burner</td>
<td>Connected to gas supply</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Box of matches</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>200 mL mixture of glucose and Benedict's solution in a beaker. Beaker labelled &quot;A&quot;.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Test tubes</td>
<td>in a test tube rack</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Test tube holder</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Safety glasses</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Thermometer 0 °C to 100 °C</td>
<td>With only one temperature scale and one degree graduations</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Water bath</td>
<td>Temperature = 40.0 °C</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>500 mL water in a 1 L beaker.</td>
<td>This beaker should be left inside the water bath</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Sponge</td>
<td></td>
</tr>
<tr>
<td>Station</td>
<td>Quantity</td>
<td>Item</td>
<td>Comments</td>
</tr>
<tr>
<td>---------</td>
<td>----------</td>
<td>-------------------------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>100 mL glass graduated cylinder</td>
<td>1 mL graduations</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>250 mL glass graduated cylinder</td>
<td>2 mL graduations</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>plastic cups</td>
<td>Prepare a set of 2 cups labelled: &quot;1&quot; and &quot;2&quot; for each student and one set as a spare</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Stick a blank label on each cup for students to write their name on.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Number the cups and obtain mass of each</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1 L plastic container filled with water</td>
<td>Labelled &quot;X&quot;</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>6&quot; test tube with 15 mL of water</td>
<td>seal top to secure water inside</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>cleaned jam jar with 125 mL water</td>
<td>seal top of jar to secure water inside</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>felt-tipped pen</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>Tray</td>
<td>Labelled: &quot;Cups for Assessment&quot;</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>sponge</td>
<td></td>
</tr>
<tr>
<td>Station</td>
<td>Quantity</td>
<td>Item</td>
<td>Comments</td>
</tr>
<tr>
<td>---------</td>
<td>----------</td>
<td>-------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>roll of red litmus paper</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>roll of blue litmus paper</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>dropper bottle</td>
<td>Labelled:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>&quot;Bottle A&quot; (contains 1/2 tsp. citric acid in water solution)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>&quot;Bottle B&quot; (contains lime water)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>&quot;Bottle C&quot; (contains water only)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>&quot;Bottle D&quot; (contains 1 tsp. citric acid in water solution)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>plastic cups (small)</td>
<td>Prepare a set of 4 cups labelled: A, B, C and D for each student and one set a spare</td>
</tr>
<tr>
<td></td>
<td></td>
<td>plastic cup</td>
<td>Labelled: &quot;Used litmus paper&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>tray</td>
<td>Labelled: &quot;Used Cups&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sponge</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX 9
Administration Guide for the Practical Test

ADMINISTRATION OF THE TEST OF SCIENCE APPARATUS SKILLS.

Stations Arrangement.
The stations should be arranged according to the following setup:

```
station 1
  [ ] supervisor B

station 2
  [ ] supplies

station 3
  [ ] supervisor A

station 4
station 5
```

Supervisors' Tasks.
This test has been designed to be administered by at least two supervisors.

Supervisor A should:
- be the main supervisor
- make the necessary announcements before, during and after the test
- observe and mark the performance of the students at station 3.

Supervisor B should:
- look after stations 1, 2, 4, and 5
- answer the questions that students may ask (questions about definitions of terms should be answered)

The Test.
This test caters for a maximum of five students per session.

Each student should be provided with a pencil and the test booklet.
The supervisors should be in the room before the beginning of the test to ensure that everything is set out properly and that all the materials required for the experiments have been placed correctly.

When everything is ready, supervisor A should allow the students to enter. Each student should be directed to a station. When the students are settled, supervisor B should hand each student the Test of Science Apparatus Skills booklet and a pencil. The booklet should be placed upside-down on the desk. When handing these out, s/he should also inform each student not to turn the booklet up until they are asked to do so.

After handing out the booklet, supervisor A should announce the following:

*You are not allowed to talk or to communicate with any student during the test. If you have any problem please raise your hand.*

*Please turn your test booklet to the second page and write your name, date of birth, class or form, and your school in the spaces provided.*

*You have five minutes reading time. During this time, you should read the test instructions, and then turn a few pages in your test booklet until you find the number of the station where you are seated and check that you have on your desk all the materials required for the experiment at your station.*

*You should not begin to write or to work on the experiment until you are asked to do so.*

*Your reading time begins now.*
After five minutes, supervisor A should announce the following:

You may start now.  (note starting time)

After five minutes, supervisor A should announce the following:

You have two minutes to complete the experiment.

When seven minutes have passed, supervisor A should announce the following:

Time is up. Please stop your experiments and close your booklets. You have one minute to clean up your station. Make sure that the station is clean and ready for the next person. Do not move to the next station until you are asked to do so.

After one minute of cleaning up, supervisor A should announce the following:

It is time to move to the next station, move in a clockwise direction

You may move now.

When the students are settled into their new stations, supervisor A should announce the following:

You are now on your second station.

You may start now.  (note starting time)

This will continue on until the students visit 5 stations. After this time, Supervisor A should announce the following:

This is the end of the test. Please hand your booklet and your pencil to me as you leave the room.

Thank you for your cooperation.

You may leave now.