Probing the Conceptions and Mental Models of Students from Diverse Educational Backgrounds in the context of a Science Centre Show on Sound

by

DEREK BRUCE FISH

Submitted in fulfillment of the academic
requirements for the degree of
PhD in Physics in the
School of Chemistry & Physics,
University of KwaZulu-Natal
Pietermaritzburg

December 2016

As the candidate's supervisor I have approved this thesis/dissertation for submission.

Signed: ______________

Name: Professor Trevor R. Anderson

Date: ________________

Abstract

This study seeks to contribute to science education in South Africa by probing the conceptions and mental models of students in the context of a science show on sound performed at a science centre. There has been a dearth of literature on the topic of sound, on science centres and shows and with research located in Africa, hence the need for this study. The subjects in this study were grade 9 students from three different school groups: urban, township and rural. The students were visiting the Unizulu Science Centre and attending a science show, "Good Vibrations", which presents concepts in sound and waves through the medium of musical instruments. Three frameworks guided this study: a four-level framework was used to investigate and classify student difficulties with sound, and to modify these in the light of data. The CRM model was used as a framework to analyse the show as an external representation and to categorise student learning. Design Based Research informed the process of modifying both the show and the probes used to evaluate it in an iterative process. The following research questions were accordingly addressed: 1) What types of probes would effectively reveal evidence of student learning about sound and any related conceptual and visual difficulties after experiencing a science centre show *on the physics of sound*? 2) How do the students from the *urban, township and rural* school groups compare, after experiencing *the sound show*, in terms of: a) General attitude; b) Conceptual and visual difficulties with respect to sound and c) Prior knowledge and learning relating to the show. 3) Building on the data obtained from addressing RQ 2, how can the show be redesigned to improve understanding and minimize conceptual and visualization difficulties with sound? 4) Building on the data obtained from addressing RQ 1, how can the probes be modified so as to be more effective and valid tools for revealing evidence of student learning and any related conceptual and visual difficulties after experiencing *the sound show*? A combination of qualitative and quantitative probes was developed and administered to the students as pre- and post-tests. The instruments were carefully validated and then inferences drawn from the student data from a statistical analysis and item analysis. The probes showed evidence of significant learning taking place during a science show with two of the three groups doubling their pre-test score in the post-test. The show worked well for urban and township students and lessons were learnt about how to make the experience more meaningful for the rural students, by paying attention to issues of language, cultural background and prior knowledge. Following modification of the show and retesting, new student data showed little change for the urban and township groups but a large impact for the rural group. A final modification saw the show presented in isiZulu to a new rural group that fared even better in the subsequent testing. The major novel achievements of this study are the following: Firstly, in addition to documenting many student difficulties from the literature in the area of sound and waves, three novel student difficulties were identified and classified, namely: that sound travels in electro-magnetic waves; that sound is unaffected by obstacles, passing right through them and that sound waves turn towards a hearer. Secondly, a thoroughly validated set of probes was developed that is available for other practitioners to use. Thirdly, improvements to the show and the probes made measurable difference to the weakest rural group. Fourthly, presenting the show in isiZulu to a rural group produced significantly better results in an English language-based evaluation. Finally, a comprehensive practical guide to developing, evaluating and improving science shows was developed for practitioners. With the paucity of science centre research, it is believed that this study will provide much-needed motivation for more resources to be invested both in further research and indeed in science centres themselves. These results are also more broadly applicable to science education in general.

Preface

The research described in this thesis was carried out from 2007 to 2016 in the Discipline of Physics, under the supervision of Associate Professor Trevor R. Anderson, Department of Chemistry, Purdue University and Senior Research Associate, School of Life Sciences, University of KwaZulu-Natal, and under the co-supervision of Associate Professor Nancy J. Pelaez, Department of Biological Sciences, Purdue University and Associate Professor Saalih Allie, Department of Physics, University of Cape Town.

The work presented in this thesis represents the original work of the author and has not been otherwise submitted in any other form for any degree or diploma to any other University. Where use has been made of the work of others, it has been duly acknowledged in the text.

Signed:

____

Derek Bruce Fish (Candidate)

____

Prof. Trevor R. Anderson (Supervisor)

____

Dr. Nancy Pelaez (Co-Supervisor)

____

A/Prof. Saalih Allie (Co-Supervisor)

Contents

List of tables

PAGE

List of figures

PAGE

Acknowledgements

- To my dear, patient, wife, Karen, who has put up with my stress and supported me all the way, believing in me even when I didn't. This wouldn't have happened without your love.
- To our precious children, Jessica and Tammy, thanks for your patience with me.
- To my parents, Stan and Barbara Fish, who stimulated my curiosity at a young age and still advise and encourage their son.
- To my supervisors, Professor Trevor Anderson, Dr. Nancy Pelaez and Dr. Saalih Allie. Thank you for persevering with my rather staggering progress and seeing me through to the end.
- To my three examiners, who provided constructive and useful feedback which improved the final thesis and will inform future publications.
- To UKZN and the NRF for the grant which covered my studies.
- To Alfred Tsipa and Marty Schwartz at Unizulu Science Centre for being part of the expert input to validation and for all your ideas and suggestions.
- To my other colleagues at Unizulu Science Centre: Reuben Nzimakwe, Mdumiseni Nxumalo, Simphiwe Mthiyane, Silindile Mthembu, Diloshni Thambaran and Zamani Ncanana with whom I have worked for so long. Thank you for covering for me when it got busy.
- To the funders of Unizulu Science Centre (too numerous to mention) who make it possible for us to tackle the giant of science education in South Africa.
- To the children of Zululand who are so hungry to learn – I trust this study will help me to serve you better.
- To the principals, teachers and students of the three schools who allowed me to study them. Thank you for assisting me.
- To Felixton College for providing bus transport to assist rural learners to get to the Science Centre for the study
- To God who has given us a living hope as an anchor for the soul, and who has guided, sustained and strengthened me through an incredibly tough project.

Chapter 1: Introduction, Aims and Research Questions

Quality science education is a key issue in South Africa (Reddy et al., 2012) as the nation's future depends on the extent to which we can educate and skill the workforce of the future (Hanushek, E. A., & Wößmann, L. 2007). The 2016/17 World Economic Forum Global Competitiveness Report has recently ranked South Africa last of the 138 countries surveyed, in terms of the quality of our maths and science education (www.weforum.org). South Africa was also ranked last in the 2015/16 report. The same report shows many areas in business where South Africa scores well, but "an inadequately educated workforce" is seen as the third most problematic factor for doing business in this country (ahead of corruption and crime). This problem has a long history: under Apartheid education (up to 1994) Black students were denied access to quality education and especially to science, which was considered not necessary or useful for them. The government of the time spent ten times as much per white student as they did per Black student. Many Black schools were without water and electricity and almost none had science laboratories or qualified science teachers (Christie 1986).

Since 1994, there does not appear to have been substantive improvement in the education system as a whole and especially in the realm of science education (Jansen 2008). While much has been made of increasing matric pass-rates, they may have been skewed by a shift away from "higher grade" subjects more than by improved education (Dempster 2004). One indicator of this is the Trends in International Mathematics and Science Study (TIMSS: www.timss.bc.edu), results of which have been published every 4 years since 1995, where South Africa has mostly been ranked last of the fifty or so nations participating. Reddy (2006) described the achievement of South African learners in TIMSS 2003 as "very low" and "reflecting the inequalities many learners are confronted by within the education system itself." The TIMSS 2015 results (released 30 November 2016) show that South Africa has made great improvement since 2003, but still lies last or second-last on the table.

Thus the results of unsuitable curricula, unqualified teachers and a lack of resources have been poor matric results and low throughput of students into tertiary science courses and science careers (Naidoo and Lewin 1998). For example, industries in the Richards Bay area of KwaZulu-Natal around the Unizulu Science Centre, the context of the present study, constantly bemoan the short supply of suitably qualified candidates for bursary programmes and jobs. Thus there is an urgent need to address these problems both in KwaZulu-Natal and throughout South Africa through studies like this one.

There have been many ventures to attempt to address this crisis, for example non-governmental organisations like: CASME: The Centre for the Advancement of Science and Mathematics Education (James et al., 2008); PSP: The Primary Science Programme (Harvey, 1999) etc., including my own and others' involvement in interactive Science Centres. South Africa has about 30 Centres of various sizes who belong to the Southern African Association of Science and Technology Centres (SAASTEC, see www.saastec.co.za). Section 2.5 presents a review of published educational research on Science Centres. What is clear from the review is that a dearth of such research has been done in Africa, despite the massive potential of science centres for impacting on the education of young people in science. Science Centres seek to assist in motivating and instructing students and teachers through science shows, small group workshops and interactive exhibits. Science shows in particular have been researched very sparsely and there is a great need to understand more about them in terms of their educational and motivational benefits, especially in my local context where young people don't always have access to high quality science education. While others have conducted studies of student learning in science shows (Sadler, 2004; Walker et al., 2011), this study is the first (to my knowledge) to attempt to apply that knowledge to improving the show and measuring that improvement. Thus there is a need to find out how these shows can be effectively evaluated and consequently improved.

Unizulu Science Centre, of which I have been director for 25 years, is one of these science centres and the context for the present study. In brief our major goals are: "To make science real, relevant and fun; to assist students to pass science; and, to encourage students to continue with science after finishing school." The Centre has a high throughput with 30 000 students visiting per year. There is some debate about the effectiveness of science centres with Bradburne (1998) referring to them as "dinosaurs and white elephants unsuited to the 21st century", but Perrson (2000) assures us that "science centres are thriving and going strong". Since that exchange some time ago, though, my personal view is that the literature has leant towards science centres being effective places of learning. This view is amply supported by the survey of 180 studies published by Garnett (2002). It is clear that to answer this question broadly, is a large scale enterprise that would involve probing both micro and macro issues. The present work, while focussing on the former, fits into this overall framework as well as overlapping with curricular issues at school level.

One of the situations that is unique to science centres in general and to the Unizulu science centre in particular is that, unlike schools who draw their students from a more limited sector of society, visitors to science centres come from more diverse backgrounds. In this regard, Reddy (2006) reported from the TIMSS study how these different backgrounds can affect student achievement in science and mathematics. Since the Unizulu Science Centre is visited by three distinct groups of students of different educational and social backgrounds, namely from urban, township and rural schools, this afforded me the opportunity to investigate in the present study any variations in student performance across these groups. More detail on these different groups is given in 3.1.2. In the present study I shall compare these groups of students with respect to their conceptual understanding and mental models of sound and their extent of learning about this phenomenon during one of our science shows presented at the centre. Whereas other studies have compared groups at different ages or levels of physics instruction (McCloskey, 1983), the present study compares the experiences of students of the same age from different socio-economic backgrounds. Such insights are considered crucial if we are to meet the diverse needs of all visitors to our science centres.

The topic of sound and waves forms part of the school curriculum at Grade 10 level in the new CAPS curriculum (see section 3.1.3). Formerly it was in Grade 11 in the NCS syllabus. Sound is one of the areas of Physics which has been least studied (Linder, 1992) with most of the attention focussing on mechanics, electricity and light (Caleon and Subramaniam, 2010). It is an area in which students have first-hand experience (through speech, hearing and music) of the concepts involved and yet is problematic to teach, with students' naïve theories often not aligning with currently held science (McCloskey 1983). In a similar study on student understanding of concepts in mechanics, McCloskey (1983) asserted the need to characterize students' naïve theories not only in mechanics but in other scientific domains as well; hence this study on sound.

To my knowledge, apart from Linder (1993), none of this research has been done in Africa, but instead mostly in countries that perform very well in Science (according to the TIMSS tests, www.timss.bc.edu) like Singapore (Caleon and Subramaniam, 2010). Thus there is a pressing need to look at student conceptions in the local Zululand area, which has some of the worst performing schools in South Africa as well as in the world. Most of the studies have also looked at University students (Linder, 1993), or students who have completed a section of instruction on sound (Hrepic et al., 2010), rather than secondary school students. The goal of the present study was to compare local Grade 9 students who have not yet formally studied sound in class. I wanted to investigate their naïve, everyday conceptions and mental models of sound *before* they were affected by input from teachers – either good or bad. This would form a useful baseline for future teaching or show design.

So, in summary, this study addresses key issues in the critical area of science education in South Africa, and attempts to fill a much needed gap in the literature, especially that between educational research and practice (Levin and O'Donnell, 1999). There has been a dearth of research into science centres and shows (especially in Africa) and this is the first study which will look not merely at learning from a show, but also at how to attempt to improve that learning. In many cases this dearth of research has resulted from the lack of suitable probes to evaluate learning. The topic of sound has not been widely researched, especially at secondary school level and in Africa. Finally a comparison of the unique experiences and needs of diverse student groups visiting science centres will contribute greatly to their effectiveness in Africa.

My goals in this study are thus to investigate the attitudes, everyday conceptions and mental models of sound held by diverse students before and after experiencing a science show. To that end I would develop and refine probes to reveal evidence of student learning and use this evidence to redesign and improve the show. Towards these goals, I addressed the following specific research questions.

Research Questions

1) What types of probes would effectively reveal evidence of student learning about sound and any related conceptual and visual difficulties after experiencing a science centre show *on the physics of sound*?

2) How do the students from the *urban, township and rural* school groups compare, after experiencing *the sound show*, in terms of:

a) General attitude;

b) Conceptual and visual difficulties with respect to sound;

c) Prior knowledge and learning relating to the show.

3) Building on the data obtained from addressing RQ 2, how can the show be redesigned to improve understanding and minimize conceptual and visualization difficulties with sound?

4) Building on the data obtained from addressing RQ 1, how can the probes be modified so as to be more effective and valid tools for revealing evidence of student learning and any related conceptual and visual difficulties after experiencing *the sound show*?

A literature survey was conducted (chapter 2) for theoretical frameworks and research on learning theories and student conceptions on sound and waves. After careful consideration of the relevant contexts (section 3.1) a science show was developed (3.2) as well as probes for evaluating student responses to the show (3.3 – 3.5). Both the show and the probes were piloted (3.4.1 and 3.4.2) and then a first (*original*) round of testing done. Data from this testing informed a thorough validation of the probes in answering RQ 1 (chapter 4) and allowed for a comparison of the student groups in answering RQ 2 (chapter 5). In the light of the data from RQ 2, the show was redesigned through a Design Based Research (DBR) process (chapter 6) in answering RQ 3. In addition the validation data from RQ 1 informed a redesign of the probes in answering RQ 4 (chapter 7). A *subsequent* round of testing with new student groups provided data for a *subsequent* validation of the new probes (chapter 8) further answering RQ 4 and a second comparison of the school groups (chapter 9), further answering RQ 2 and informing RQ 3. A *final* redesign of the show saw the show (but not the probes) translated into isiZulu (10.1 and 10.2) giving further input towards answering RQ 3. A *final* test of this show with one new student group (10.3 and 10.4) provided one more set of data towards answering RQ 2. Answers to the RQs were summarised in chapter 11. A detailed layout of how these questions are addressed and in which sections, is given in tables 3.2, 3.3 and 3.4 in section 3.3. A similar layout in terms of a DBR framework is given in table 2.3.

Chapter 2: Literature review: Theoretical Framework

Science Centres around the world present science shows to millions of people annually, yet there is little reflection on or evaluation of these shows. What does occur is usually of a "visitor-survey" nature, and aims more to provide motivation for funders than to research deeply what visitors gain from these shows. In South Africa in particular, science centres are in their infancy and this feedback is desperately needed to assist practitioners towards best practice. The literature reviewed in this chapter was carefully chosen for its applicability to this project and to science centres in general. It is intended to provide a theoretical framework for this study, outlining those theories which guide and frame the study by focusing on published research. As I am investigating conceptual difficulties and learning (RQ 2), it begins with an overview of: learning theories and models (2.1), conceptual understanding and conceptual change (2.2) and visualization and mental models in science (2.3). Section 2.3 presents the CRM model of Schönborn and Anderson (2009) as a way to frame development and analysis of the show and the evaluation probes. In order to characterize and contrast the very different school groups in the study, I must consider first: the influence of culture and background in learning in science (2.4). As the study takes place in a Science Centre, rather than a school, I consider from the literature material on the role and effectiveness of Science Centres and science shows (2.5). Then, as a starting point for my study I conduct an extensive literature review of student conceptions of sound (2.6). This culminates in the formulation of a detailed student difficulties table (table 2.2 in section 2.6) which guides the rest of the study and forms the basis of the probe design and evaluation. The table rates the difficulties according to the 4-level framework of Grayson, Anderson and Crossley (2001) and provides propositional knowledge statements for each difficulty. Finally, I consider Design Based Research (Trujillo, Anderson and Pelaez (2016) and other references) as an effective framework for efforts to remediate both the show and the probes (2.7.1) and apply it as a framework to shape the subsequent sections of this thesis (2.7.2).

In summary then, three different theoretical frameworks are employed for this study:

Table 2.1: Theoretical frameworks in this thesis and details of their application:

Theoretical Framework	Main Reference	Introduced in section:	Used in this thesis to frame:	Applied mainly in sections:
CRM Model	Schönborn and Anderson (2009)	2.3 (Figure 2.1)	Development of ER (science show) and probes	3.2 – 3.5 (tables 3.9 & 3.10); Appendix a Chapters 4 - 11
Four-level framework	Grayson, Anderson and Crossley (2001)	2.6 (Figure 2.2)	(Re) Classification of student difficulties	2.6 and 5.2 (tables 2.2 & 5.2)
Design Based Research	Trujillo, Anderson and Pelaez (2016)	2.7 (Figure 2.5)	Main (re) design process of show and probes	Chapters 2 – 11 (tables 2.3 & 11.2)

2.1 Learning theories and models framing this study

Most of what has been written about different theories of learning applies to more formal learning in the classroom. It must be stressed that this is a study of learning that takes place (or doesn't) in an informal learning environment- the science centre. I shall be looking at the effect of a single interaction (see section 3.3) with students from different backgrounds (see section 3.1.2) in the context of a Science Centre (see section 3.1.1) rather than a school. While some tenets of these learning theories will apply to a science show, most apply to more sustained instruction given over a longer period.

My experience over 20 years as teacher and Science Centre Director has shown me that students from different backgrounds have a very different experience of a Science Centre visit: this is what led me to do this study. It would be naïve and incorrect to assume that one programme would suit and benefit all visitors in the same way. Generally speaking, students are products of their backgrounds, and the way they behave and interact during a visit reflects this. Constructivist learning theory seems ideally suited to this situation as it describes learning in terms of an interaction between knowledge the students already have and new scientific knowledge they will encounter (Pines 1986). Constructivism stresses that people create their own meaning of the world through experiencing it first-hand in order to gain and indeed to test new knowledge. Learning is an active process whereby students make sense of new information through their own experiences (Driver 1989). Prior knowledge and ability is critical as it affects (either positively or negatively) whether new concepts will be successfully learnt and mastered (Von Aufschnaiter and Von Aufschnaiter, 2007).

Constructivism arose as a counter to the learning theories of the 1960's and 1970's, which largely ignored the contexts of the students, treating them as "empty vessels" to be filled with knowledge. Even today though, von Glasersfeld (2010) and others have confirmed that although well supported empirically, constructivism has not become a generally accepted theory of knowledge. He borrows from Freudian terminology to talk about the "stickiness of beliefs" which prevents people letting go of long cherished ideas. In addition, he points to the "thirst for certainty" in both students and teachers. Matthews (1993) attempted to show the problems constructivism faced in the light of Aristotle's empiricist epistemology and Kirschner et al. (2006) argued that the minimal guidance of constructivism does not work *at all* in: "The Failure of Constructivist, Discovery, Problem-Based, Experiential and Inquiry-Based Teaching.

All this is naturally of great importance for Science Centres, for whom Discovery Learning is their central basis for existence and many of which are based on constructivist principles (Sjoberg, 2007). Indeed in my Centre, one of my mottos is the ancient Chinese proverb: "Tell me, I forget, Show me, I remember, Involve me, I understand!" Science Centres in many ways arose because of the traditional teaching in science classrooms which was felt to be boring, uninspiring and ineffective. (See section 2.2). There are different aspects to a science centre visit though, with various levels of student involvement which are more fully described in sections 3.1.1 and 3.3. The three main activities of most centres like ours are: science shows (where an expert demonstrates to groups of 80 – 160 students), small group workshops (where 6 – 8 students take part in some small group activity) and interactive exhibits (where facilitators assist students individually or in pairs to explore exhibits). Thus the centre provides activities focused on: large groups, small groups and individuals. In terms of the above discussion, Science Shows fall more under Kirschner et al.'s (2006) "guided instruction" while exploring the exhibits is more of a discovery learning activity (with minimal guidance from facilitators). The small group workshops involve a combination of these two extremes.

For reasons of brevity, this study will focus on science *shows*, or one show in particular, and how different school groups respond to it. I have been performing science shows for the last 20 years and am convinced that a good show does not have to adopt a "tabula rasa" approach to the children. Much can be done with audience response (including via remote clickers), use of volunteers, question and answer, participative experiments, demonstrations which ask for predictions etc. to make the show highly interactive. Nevertheless one cannot truly adopt a constructivist approach in presenting the show, unless one has taken the time and trouble to gauge the students' abilities, attitudes and prior knowledge. These issues are addressed in the present study.

In addition to the mixed modes of delivery used in a SC, visiting students are drawn from vastly different environments. The three different contexts from which my study groups will be drawn are described in detail in section 3.1.2. My experience has shown that this background affects the students in every part of a Science Centre visit. It is more overt in workshops and with exhibits as students are more passive during the shows, which is why I wanted to investigate shows specifically. Unfortunately, the very nature of a science centre visit often militates against an approach which takes the students' prior knowledge and backgrounds into account. Time is limited, and there is usually no time to prepare for the visit beforehand, or evaluate it afterwards. Many science centres in foreign countries do pre-visit preparation activities (many over the internet), but these are impossible for us given the remote, rural nature of our schools. Evaluation so far has been of a very limited "visitor-survey" nature usually conducted only with the teacher, which does not look deeply into what students actually experienced. It is hoped that this study will provide useful information both on prior knowledge and abilities, and experience of the visit, which will be generalisable to future visitors. This can then provide a background allowing us to tailor presentations for particular groups.

Considering the wider perspective of education in general, the learning theories in this section are also very important. While I have been running the SC, there have been several curriculum-related reforms aimed to democratize education and to eliminate apartheid inequalities- from the old National Senior Certificate (NSC), through Outcomes Based Education (OBE), then from 2008 the National Curriculum Statement (NCS) and finally from 2012 to the present date the Curriculum and Assessment Policy Statement (CAPS). OBE in particular, was an attempt to adopt a more constructivist approach and to allow for more student-centered, discovery learning and less top-down teaching. For many reasons OBE failed (Jansen, 1998, lists ten different reasons) as a system unsuitable for South Africa, perhaps largely because of large classes, lack of resources and un- or under-qualified teachers.

With all the changes in curriculum there has been little tangible evidence of change in the teaching and learning of science. While increases in matric pass-rates have been heralded, many researchers (e.g. Dempster, 2012 and Jansen, 2012) have attributed this more to lowered standards than raised capability. From my regular visits to Zululand schools over the last 20 years, I can only agree with these authors. International benchmarks (like the TIMSS tests: www.timss.bc.edu) have shown no improvement and even some worsening of our standards. It is not yet clear whether the CAPS curriculum will fare any better. In theory, CAPS does at least stress the importance of practical investigations: "In Grade 9, students are expected to carry out their own investigations to expand on the concepts or knowledge to which they have been introduced and to deepen their understanding of the subject matter." (CAPS Natural Sciences Document: July 2011)

Sadly the schools around the Unizulu Science Centre are still mostly teaching science without any experiments, or even teacher-led demonstrations. Discovery learning in natural sciences is not possible without at least some basic equipment which students can use to make sense of new information through their own experiences (Driver 1989). The challenge in South Africa is not only to find the learning theory that will inform best practice, but to provide the resources (both human and physical) to transform theory into reality. Science Centres have always seen themselves as being part of that process.

2.2 Conceptual Understanding and Conceptual Change Theory

We shall accept the assertion that students' prior knowledge and abilities must be taken into account before effective learning can occur. Students often hold their own conceptions which are different from Physics conceptions, but are nevertheless tenaciously maintained, even though they may hinder the learning processes of sound scientific conceptions (Chi and Roscoe 2002). These are variously referred to as misconceptions, preconceptions, naïve conceptions or alternate conceptions. This has developed into a complete field of study: Conceptual Change Theory, which has grown over the last three decades (Duit and Treagust 2003). I shall refer to them simply as student *difficulties* (Grayson, Anderson and Crossley 2001) which in the present study were pre-instructional since none of the students had received any prior formal instruction in the phenomenon of sound- only every-day, experiential learning. Such difficulties are especially prevalent in Physics, where students' personal interactions with the natural world cause them to form their own notions of concepts like gravity, pressure, speed and colour. I have chosen sound and music for this study, an area where students form their own conceptions from childhood of volume, pitch, rhythm, waves, vibrations and a host of other concepts.

The challenge then for a science teacher (lecturer or science centre staff member) is how to change these concepts into more scientifically acceptable ones: or "conceptual change" theory. Piaget introduced the terms assimilation and accommodation to describe how this can happen. Assimilation has become most closely connected to constructivism and involves incorporating the new experience into your inner world by fitting it into existing mental schemes, or "pigeon-holes." This is less stressful for the student, but only works if the existing schemes are able to accommodate the new concepts.

Accommodation involves our internal world accommodating itself to the new concepts by changing our existing mental schemes. This is obviously more stressful for students and can be met with great resistance – especially where existing mental schemes and concepts are strongly believed and held. The aim is for the student's mental schemes to be adjusted so that he or she can make sense of the external evidence he or she is exposed to. I believe that many of the problems we have in Science Education in South Africa arise because no experiments are performed at schools. In the absence of empirical evidence, students are less likely to change their mental schemes just because a teacher says so. Again, I believe the great advantage of science centres is that they present first-hand, physical experience to students during the shows, workshops and exhibit exploration times.

In the light of the above, the first challenge of this study is to probe student understanding in order to identify prevalent student difficulties which exist in the area of sound, waves and music. A literature review is presented in the section 2.6 and the difficulties (along with some which I have identified over the years) categorized using the four-level framework of Grayson, Anderson and Crossley (2001). Details of this categorization are given in table 2.2 in section 2.6. Probes were then designed to specifically test for evidence of these difficulties, and whether there was any change after viewing the show. It is also anticipated that new, unanticipated difficulties may come to light through this process. It is of particular interest to see whether these student difficulties are common to all of the three groups in the study, or peculiar to one or two of them.

Conceptual *Change* theory focuses not so much on the conceptions themselves, but on how they can be effectively changed. There are various conflicting approaches to describing change processes depending on the theoretical stance of researchers. Thus, di Sessa (2002) from a “knowledge in pieces” perspective describes conceptual change as “the reorganization of diverse kinds of knowledge into complex systems in students’ minds”. This dynamic perspective of conceptions, namely, in-the-moment (or “on the fly” as in Cooke and Reedin, 1994) constructions that are strongly context dependent can be viewed as being in contrast to the more static and unitary “misconceptions” perspective (McCloskey 1983) which leads to conceptual change being described by the “repair of misconceptions” (Chi and Roscoe 2002). In the latter view students start with their naïve conceptions relative to the canonical view of Physics, identify where they are incorrect and repair them. Whether one takes the former view or the latter the role of the teacher is critical and there has been much debate about exactly what this should be. Again the science centre context is very different, in that one is a presenter rather than a teacher, and there is no opportunity for sustained input into this “repair” process. In addition, in the present study, students arrived for the show with only conceptions acquired informally from life experiences rather than formal instruction, as they were Grade 9’s who had not yet covered sound at school.

Posner (1982) provided four useful conditions which encourage students to take on the accommodation required for conceptual change, (if following the static misconceptions approach):

1. Dissatisfaction with their existing conceptions and a realization of their inconsistencies
2. A new conception which is intelligible and makes sense
3. A new conception which appears initially plausible and makes more sense than the old concept
4. A new conception which suggests the possibility of fruitful new areas of inquiry

It is important to note that these conditions do not guarantee conceptual change, nor are they intended as a teaching method, but rather as a description of the process through which change can occur. Often very special strategies are required to facilitate conceptual change and these can be very different for different concepts and contexts. Niebert et al. (2012) stressed that merely linking the science to everyday life does not guarantee fruitful science communication.
While it is necessary for the above four stages to happen for conceptual change to occur, it does not imply that it will always occur. Duit (2003) has provided many examples of conceptual change failing to occur. Taking into account the very limited teaching opportunity during a single science show, it will be interesting to investigate what sort of learning (if any) is evident in the students after the show. The above conditions can be very useful guidelines for Science Centres, though, especially in the area of show design.

2.3 Visualization and Mental Models in science

In this study, students are exposed to a brief science intervention in terms of the science show entitled "Good Vibrations" which uses various models to attempt to explain phenomena in sound. Different from standard lessons or textbooks, the elements of this show (or models) can be considered as External Representations (ER's), which attempt to explain physical phenomena in sound (like frequency, volume, noise etc.) through various graphical, pictorial and visual means. Gilbert et al. (1998) presented four representational modes in which these models may be expressed: material, visual, verbal or symbolic. Section 3.2 gives a detailed description of the show including its rationale and design, and describes the many different multimedia elements which are combined in the show, including all four of the modes mentioned above. The show attempts in some ways to make up for the very deprived science education most of our students receive (see section 3.1.2) in physics, specifically to do with the phenomenon of sound.

This study seeks to compare and contrast the response of students from three different backgrounds to the various ER's or models presented in the show and the various factors affecting their response, including their prior knowledge and experience, their social backgrounds and the affective domain including motivation and attitude to science and science learning. Gilbert et al. (1998) stress the fact that a student's *mental* model thus formed is only accessible when it becomes an *expressed* model (in action speech or writing) and will commonly differ from the *consensus* model which is currently accepted by the scientific community. The presenter is searching for a *teaching* model which aids students to understand a given consensus model. An initial study by Schönborn et al., (2002) identified three factors which could cause student difficulties with the interpretation of ER's (in Biochemistry, but transferable to Physics). These were students' reasoning ability (R), their understanding of the concepts of relevance to the ER (C) and the nature of the ER itself and how well it represents the phenomena, or the mode (M). These are very applicable to my situation and could be helpful in determining whether the ERs presented in my science show, were successful.

In a follow up study, Schönborn and Anderson (2009) extended this model to look at interactions between pairs of factors, and even all three, the "C-R-M Model". The reasoning factor, R, is now seen in terms of the student's ability to reason with both the ER itself (this intersection is labeled R-M) and with his or her existing pre-conceptions (R-C) of the subject matter which is relevant to the ER (C-M). This linking is in line with the "active learning" of the constructivist paradigm which is outlined in 2.1, in that it recognizes the prior knowledge and experiences of the student. It must be stressed that conceptual difficulties (C) will be directly related to the context (sound in this study) whereas reasoning difficulties (R) can be context independent and can occur between contexts. For my study a conceptual difficulty could be the confusion between pitch and volume. An example of a reasoning difficulty (in this case a visualization difficulty) would be students confusing the representation with the reality in: vector diagrams, velocity-time graphs, electric circuit diagrams and drawings of sound propagation (Van Fraassen, 2010).

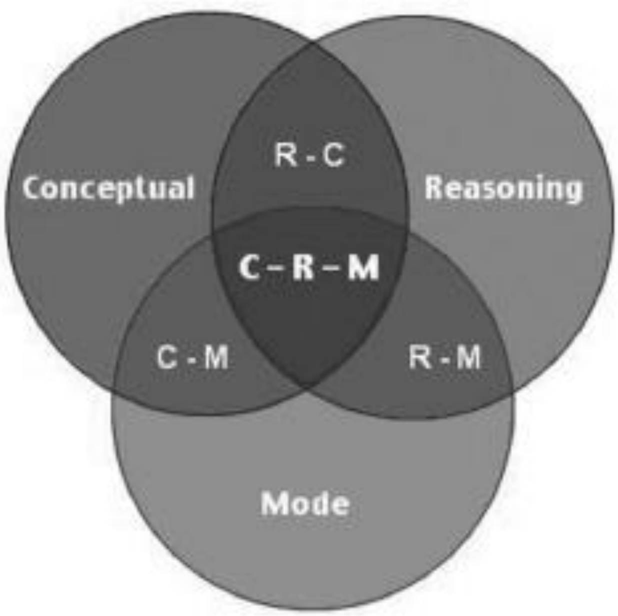


Figure 2.1: CRM Model of Schönborn and Anderson (2009)

This allows one to consider the success (or failure) of the ER's in terms of students' ability to make sense of, visualize and learn from the ER's they have experienced (C-R-M) in terms of the criteria suggested at the end of the study (Schönborn and Anderson 2009). The model stresses the importance of the interplay between factors affecting students' ability to learn from the ER and how the model can be used to determine which factors most influenced the students' learning and interpretation.

Part of Schönborn and Anderson's (2009) conclusion is that: "the model has a generic application to all types of ERs in science, including not only static representations but also dynamic, animated, and multimedia representations." This seems to indicate to me that it is applicable to the elements presented in a science show like "Good Vibrations." Hence in a study like ours, the mode, M, would represent the elements of the science show (section 3.2 and appendix a), C the concepts and knowledge of importance for understanding the topic of sound (also in section 3.2 and appendix a), and R the different cognitive, visual and interpretative skills the students would require to master this subject area (discussed with development of probes in section 3.4).

This model thus provides a useful framework which, as illustrated in Anderson et al. (2013), can be used to guide the selection of research questions, identification of student difficulties, design and testing of probes, gathering and processing of student data and formulation of conclusions and implications.

2.4 The influence of culture and educational background on learning in science

The constructivist learning theory stresses that learning is not the same experience for each student and that they make sense of things with reference to their own experiences. Section 3.1.2 details the vastly disparate contexts from which the three student groups come. South Africa has the highest GINI coefficient in the world (World Bank) indicating a vast gap in wealth distribution, which is mirrored by a similar gap in distribution of educational resources. What has been striking about SA's results in the TIMSS tests are not just that we achieved the lowest marks, but also that our marks showed by far the highest spread of any country. In South Africa there are dramatic differences in provision of resources and qualified teachers which are obviously major factors. Furthermore, because science is taught in English, second or third language speakers are immediately disadvantaged. In addition to that, though, a student's home environment makes a big difference to how he or she learns science (Lee, 2003). It is important, therefore, for me to consider the influence that culture and background has on learning in science. In the present study, the rural group is schooled in circumstances very different from the urban group, and I am interested to see the effect this has on various indicators.

Science achievement is controversial across the globe as “science assessments often contain a high level of linguistically and culturally dependent content that can exacerbate the persistent gaps in science achievement and professions” (Penfield R and Lee 2009). Growing up in different language and cultural settings clearly gives students different starting points for learning science. Solano-Flores and Nelson-Barber (2001) identified five areas in which culture will affect a student’s performance in science: the student’s epistemology, their language ability, their cultural worldview, cultural communication and socialization styles and their life context and values. Gilbert et al. (1998) suggest that scientific explanations can often be seen as inappropriate, or even as non-explanations in other contexts as they are drawn only from the pool available within orthodox Western science.

Lee (2003) stresses that students from different language and cultural backgrounds come to school with their own prior knowledge assimilated from their language, culture, home and community. This is often quite different from that of traditional “Western” science curricula, which can lead to a disconnection. An example of such a disconnection in KwaZulu-Natal is the fact that most rural students would consult a traditional healer when sick, whereas most urban students would see a doctor or visit a hospital. This can cause a rift between the “primary discourse” which occurs at home and the “secondary discourse” in the school (Cobern and Aikenhead, 1998). Another common example experienced by local students is community attitudes towards lightning as a mystical or magical phenomenon, contrasted with the electrostatics they would learn at school.

Lee (2003) also highlights issues around scientific inquiry. The new SA Curriculum and Policy Statement (CAPS) stresses the importance of inquiry in its opening statement on: “What is natural science?” stating that: “To be accepted as science, certain methods of inquiry are generally used.” It continues to stress inquiry as an important part of “acquiring knowledge.” As stated by Lee (2003), “Recent efforts to provide culturally congruent science instruction indicate that when linguistic and cultural experiences are used as intellectual resources, students from diverse backgrounds are able to engage in scientific practices and show significant achievement gains.”

But, inquiry is difficult for all students but even more so for students from cultures which do not encourage questioning and inquiry – especially of authority figures (like teachers). When I had just commenced working in the SC, I asked a student if he was allowed to ask questions in class. His reply has always haunted me: “We are, but if the teacher does not know the answer, he hits us”! Thus in many rural communities especially, respect for authority figures may be valued above reasoning or questioning. Thus students may get no opportunity to practise inquiry or questioning at home. This may well lead to their being satisfied with the first answer given to them by a teacher, rather than wanting to probe it more deeply or examine it further. This can come out in assessments where students may give very brief answers because their cultures do not encourage extended or detailed responses (Solano-Flores and Nelson Barber, 2001). Teachers may perceive these students as lacking complete or comprehensive knowledge.

For rural students (or any students from different cultural groups), their worlds of home and school are so different that the “assimilation” described by Piaget (see 2.2) is almost impossible. They have to succeed in the much more difficult and stressful process of “accommodation”, adapting their internal worlds to accommodate new knowledge by changing their mental schemes. These transitions have also been described as “border-crossing” (Giroux, 1992; Aikenhead and Jegede, 1999) and students’ success in science often depends on how well they can cross these borders.

Phelan *et al*. (1991) devised a model to describe how students may move from one world to another, identifying four patterns of border-crossings:

1. congruent worlds which support smooth crossings;
2. different worlds requiring crossings to be managed;
3. different worlds leading to hazardous crossings; and,
4. highly discordant worlds which cause students to resist crossings, making them almost impossible.

The “different worlds” of the students in this study are described in 3.1.2. and need to be carefully considered when evaluating the students taking part. Phelan’s model along with Posner’s (1982) conditions (2.2) encouraging accommodation may form a useful basis for considering how to improve science centre (and teaching) activities to best suit different groups. This will be considered in section 6.5. One of the reasons for choosing a science show which teaches sound through the example of musical instruments, is that much of the subject material will be familiar to rural students (use of traditional instruments, music, rhythm etc.) which I hope will make the gap between the two worlds as small as possible, and the “border-crossing” as painless as possible.

Finally, a word needs to be said about different approaches to understanding diagrams and representations, and producing drawings, across different cultures. “Good Vibrations” is full of different types of representations which students are expected to assimilate and understand. These include photographs, drawings, annotated diagrams, animations and oscilloscope traces. In addition, students were asked to provide drawings as further explanations in the questionnaire after the show. Deregowski (1989) showed that different cultural groups may use different skills to perform the same perceptual tasks with regard to diagrams. In understanding diagrams, this can cause failure to recognize a picture as a representation or even failure to recognize a pictured object at all. He cites experimental work done on “remote” populations which is relevant to our rural students. Failure to *understand* diagrams will naturally also affect one’s ability (and confidence) to *produce* diagrams. He investigates differences between “pictorial” and “non-pictorial” cultures and the difficulties they encounter with drawings. In addition, Segall et al. (1966) studied the different way in which visual illusions are experienced by people from different cultures, finding significant differences. I have observed this when anecdotally discussing optical illusions with students from different cultural groups. While this is an important area for consideration it is too vast to be considered in this thesis. Furthermore, 90 % of our students were Zulu-speaking and shared a common cultural background, so it is perhaps not as relevant a factor as issues like teacher quality, language and resource provision in the schools.

2.5 The role and effectiveness of Science Centres and Science Shows

Interactive Science Centres are quite a recent phenomenon, although it is generally agreed that the movement started in 1969 with the simultaneous advent of the Exploratorium in San Francisco and the Ontario Science Centre in Toronto. Prior to this there had been interactive exhibits in various Science museums around the world since the 1930’s, like some of those I have had the privilege to visit: London Science Museum, Palace of Discovery (Paris) and Deutsches Museum (Munich) (Stocklmayer et al. 2001). Like Prof. Mike Gore, one of the authors in the previous citation, I visited the Exploratorium (in 1994) and it had a profound effect on me, inspiring me to realize that something similar would be ideal for South Africa’s needs.

There are presently over 2500 Science Centres and Science Museums around the world, which are visited by at least 300 million visitors annually (figures from Association of Science and Technology

Centres, ASTC). South Africa is still very much at a developmental stage in terms of Science Centres, with about 30 of varying sizes existing around the country. Nevertheless South Africa hosted the most recent Science Centre World Congress in September 2011 in Cape Town. Science Centres in South Africa are affiliated to SAASTEC, the Southern African Association of Science and Technology Centres, which has as its mission:

> "To contribute to the improvement of life of Southern African nations by improving scientific knowledge and skills through the utilization of interactive living science and technology discovery centres (S and T centres)." (SAASTEC website: www.saastec.co.za)

Directors of science centres are continually asked to justify the existence and educational value of their centres, especially in the light of the constant fundraising they must do to stay open. The problem with this is that academic research is often swamped by "evaluations" which must prove the Centres valuable in order to justify their ongoing existence. Nevertheless there has been a growing body of scholarly articles recently looking at the value of Science Centres (or Museums) in the areas of cognitive learning, affective learning and motivation, and other forms of learning, which have informed my study.

In a study that involved recording visitors' conversations around an exhibit, Allen (2002) found indications of learning in terms of the quality of the content of the conversations. Visitors appeared to engage in conceptual conversation at at least a third of the exhibits. In another study Anderson (2000) looked at children's understanding of the concepts involved with electricity and magnetism during a science centre visit. They found that students not only learned facts but actually incorporated these into their understanding and models on the subject. I shall be looking for similar evidence of learning from the visit being incorporated into students' mental models and visualizations. Barriault (2001) identified eight discrete learning behaviours which she further sorted into three categories which define the depth of the learning experience in Science Centres: Initiation, transition and breakthrough behaviours.

In their book on the subject of science centres, Falk and Dierking (1992) synthesized results from their own extensive research studies into visitor activities, motivation and learning. In particular their findings stress that any museum visit involves an interaction between three contexts: personal, social and physical, which allow me to view the visit from the visitor's perspective. I have incorporated this into my study and will look at personal and social contexts in 3.1.2 and the physical context in 3.1.1 and 3.3. Stocklmayer and Gilbert (2002) conclude that the *experience* in a science centre is everything and the *engagement* of the visitor is key, more than merely pushing for greater understanding of the science. They introduce the idea of "remindings" – links in the memory to prior experience which facilitate a reworking of the idea behind the experiment or demonstration. This was taken into account in designing a sound show around a familiar subject (music) and using many local instruments which would be familiar to the students.

In a subsequent book, Falk and Dierking (2000) devote a chapter to "Documenting learning from Museums" but start with the caveat: "That people learn from museums is easy to state, harder to prove" (pg. 149). They further continue (pg. 174) that "Over the years providing compelling evidence for learning from museums has proved challenging, not because the learning did not exist, but rather because … researchers …historically asked the wrong questions and searched for evidence of learning using flawed methodologies." The authors also expand their model of three contexts, adding sub-contexts to identify eight key factors that influence learning in museums:

- Personal Context: Motivation and expectations; prior knowledge, interests and beliefs; choice and control
- Socio-cultural Context: Within group socio-cultural mediation; facilitated mediation by others
- Physical Context: Advance organizers and orientation; design; reinforcing events and experiences outside the museum

Researching all these factors would be beyond the scope of this study.

Finally, ASTC commissioned a study by Garnett (2002) which reviewed the impact of Science Centres across three continents namely North America, Europe and Australasia. She looked at 180 studies, finding the majority of these to be involved with science learning. On the whole Science Centres were seen to have a positive impact on their visitors in terms of science learning and in many other areas. Again, none of these studies was done in Africa, so it is essential that I investigate my local situation to add to this body of knowledge.

Although beyond the scope of this study, it is appropriate to mention that there is a growing body of literature concerning the affective learning which goes on at Science Centres and the role of emotion, although this has been focused more on studies in zoos and aquaria. Nonetheless Falk (2009) found a relationship between emotional arousal on a visit and positive changes in long term cognition, attitudes and behaviours. Salmi (2003) conducted a study at Heureka, the Finnish Science Centre in Helsinki, which indicated that Centres can serve as great motivators and that Centres can have a strong impact on academic career choices. Most studies stress the difficulty in causality that exists i.e. do students perform better because they visit a Centre, or do they visit a Centre because they perform better and are more diligent and dedicated? In my study visitors were bussed into the Centre rather than visiting of their own free choice.

The majority of Science Centre studies in the literature focus on the effect of exhibits. As this is beyond the scope of this thesis (our Centre has over 250 interactive exhibits, and studies have been done on the effectiveness of *single* exhibits.) I shall be focusing on the effect of a science show on student understanding, learning and visualization. Science Shows have been popular for almost 200 years since Michael Faraday began the annual "Christmas Lectures" at the Royal Institution in the 1820's (These lectures are still continuing and we sent a group of school students to attend them in 2009). These are based on the recognition that live demonstrations are essential in communicating science. A caveat must be sounded, though, as Crouch et al (2004) found that students who passively observe demonstrations understand the underlying concepts no better than students who do not see the demonstration at all. Learning can be enhanced by increasing student engagement: giving students the opportunity to discuss and predict the outcome before seeing it.

The literature on Science Shows is more limited, and I have confined myself to a few studies at science centres which I have visited, or where I have had personal contact with the presenters, or seen the shows myself. There are two interesting locally-based (or at least partially locally-based) studies that warrant mention. Walker (2011) conducted an in-depth study on an HIV-Aids Show (which I assisted in writing and presenting) as his doctoral thesis. He looked at almost 700 students and found significant changes in 9 out of 15 HIV-related intentions (i.e. intentions to change behaviour, use condoms, say no etc.) measured before and after the show. This was an inspiration to me to conduct a similar investigation with pre-test and post-test instruments. In the second study, Sadler (2004) gathered both quantitative and qualitative data from students in the UK and South Africa based on the Science Show: "Music to your ears" which covered the science of sound and waves, much like "Good Vibrations." She developed five categories, namely curiosity (C), human (H), analogy (A), mechanics (M) and phenomena (P) to describe the essence of a demonstration and used these to determine which type of demonstrations were the most memorable in the short and the long term. She found remarkable recall of demonstrations even after a 30-month period and evidence of students incorporating some of their learning into their general mental models. While this was interesting and dealt (like my study) with a music show, her study (unlike mine) was longitudinal in nature.

Shaw (2011) found that exposure to the Bristol Chemlabs outreach programme had a direct and measurable effect on enrolments of students at Bristol University. This programme includes a science show entitled "A pollutants tale" which was also performed to audiences (including myself) in South Africa at the Grahamstown Science Festival. A study by Gouniyal (2007) conducted (like Sadler's) at

Techniquest in Wales, revealed “the tremendous potential of theatre as a rich and versatile medium with the ’fun factor‘ needed for engagement, offering cognitive, social and affective gains and fostering learning dispositions, interest and curiosity.” I believe that “Good Vibrations” replicates this approach as a piece of science theatre, and intend to try to measure its effects in this study. Caleon (2005) used a multiple choice pre-test and post-test before and after a cryogenics show presented at the Singapore Science Centre. They found significant knowledge gains amongst the participants, as well as expressed enjoyment of science and increased interest in wanting to pursue science careers. To my knowledge no study has documented attempts to improve a show and measure the effect of that improvement.

Despite some advancement in trying to link science instruction more closely with the worlds of students in South Africa, we still have a long way to go. Lee et al. (2007) list six categories of challenges that still face us in this area. There is still a distinct lack of literature available on the impact of science centres and science shows in South Africa, and especially in the context of rural, non-English speaking audiences. I am convinced that this is an area where there is tremendous potential for effective engagement- hence the goal of this project is to add to the body of existing knowledge.

2.6 Review of Research into Student Conceptions of Sound

Most of the literature on this topic starts with the comment that little work has been done on students’ conceptions of sound (Eshach and Schwartz, 2006;Witmann, 2003), or that even in higher level Physics courses, sound is often treated superficially as a simple example of wave-physics (Linder, 1992). In addition,

> “Empirical studies on students’ conceptual understanding of physics concepts have concentrated on the areas of mechanics, electricity, and magnetism. There was less attention paid to waves” (Caleon and Subramaniam, 2010).

The studies all continue to assert that sound is *not* a straightforward topic, based on findings of clear gaps in students’ conceptual understanding of the subject material (Hrepic *et al*., 2010). This review was undertaken using Pfundt and Duit’s (1988) Database of student and teacher conceptions and science education, and further searches, which revealed only a few relevant papers.

Linder (1992), on summarising the findings of Linder and Erickson (1989), reported the following types of understanding of sound expressed by students: Firstly, sound is sometimes seen as an entity carried by individual molecules as they move through a medium. This entity is transferred from one molecule to another through the medium. For others, sound is a traveling bounded substance with impetus, usually represented in the form of flowing air. Another understanding was that sound is a bounded substance in the form of some traveling pattern. Finally, for many students, sound is linked to the concept of waves as part of a mathematical physics modeling system (but students' wave conceptualization is often divorced from their conceptualization of sound). Linder summarised his findings as follows:

> "… whereas in the conceptualizations based upon the microscopic perspective, sound was depicted as an entity carried or transferred by molecules with no inherent action of its own, in the macroscopic perspective sound was depicted as a substance; that is, a continuity of form that was associated with a moving "force"; a continuity of form that had an inherent action of its own." (Linder and Erickson, 1989, p. 496)

Hrepic et al. (2010) expanded on Linder's idea of the "entity model" with the notion of blend models. He identified the entity model as the usual starting point for naïve students (as would be expected in the present study), and the wave model as the scientifically accepted, or community consensus model. In addition to these two main models he enumerated models which were conceptual blends of these two, incorporating ideas from both of them. His studies showed that most students begin with an entity model of sound, but after instruction move closer to the wave model, but often finishing somewhere in between in a blend model. As my study is not a longitudinal one, this development over time could not be tested, but I did find that different school-groups tended generally to display more evidence of one particular model. Linder (1992) also discusses at length the confusion caused by representing longitudinal sound waves with transverse sinusoidal waves, and the problems associated with the "water-wave analogy", which is inappropriate at best, incorrect and profoundly misleading . From my very first pilot tests this factor came through very strongly, so was worth a detailed follow-up. He stresses later (Linder, 1993) that students' conceptions can be classified into two perspectives: a mental model (inside the head)-based perspective and an experientially-based one. My experience has shown that sophisticated, urban students tend towards more abstract mental model perspectives, while rural students living simpler lives often maintain a perspective based on their life experiences. This will be tested especially in the written answers and drawings.

While cautioning that experiments can cause further confusion if not handled correctly (like the slinky-spring and tuning-fork), in the following quotation Linder nevertheless affirms the importance of practical demonstrations and experiments- confirming the vital role that interactive Science Centres play, especially when schools do not perform experiments.

> " …even seemingly straightforward topics in Physics require more than a chalk-and-talk presentation of formulae, that much more introductory physics teaching needs to be based upon observationally based calculations and conceptual exploration. We need to teach the real thing …" (Linder , 1992, p. 263)

This is supported by Wittman et al. (2003) who stated that:

> "Based on interviews and student responses to written examination questions, we have found that lecture instruction alone (with the associated homework problems being discussed in a traditional recitation setting) had little effect on student understanding of the basic wave properties of sound propagation."

As South African schools (and especially Zululand schools) seldom or never perform experiments, this is particularly relevant. Knowing that teaching alone (especially weak teaching as found locally) has little effect on conceptual understanding; I was motivated to see if a science show (lecture demonstration) could make a difference.

In a later study, Linder (1993) identified three conceptualizations of the factors which university students' believe affect the speed of sound, namely: a *sound-resistance* factor based on the physical size or density of the molecules; a *separation* factor based on molecular separation as a function of medium density and a *compressibility* factor based on a (confused) understanding of the elasticity of the medium. Wittman *et al*. (2003) found similar problems with conceptualizations of wave-speed, whereby students believed that the force used to create a wave would determine its speed. His students also had problems with wave interactions, with the majority believing that waves collide when meeting each other and that superposition occurs only at a single point. As I was focusing on musical instruments rather than wave-speed per se, this was less relevant, but nevertheless worth taking note of.

In other studies Watt and Russel (1990) found that students perceived sound as an invisible object with dimensions which required some space in order to move, echoing Linder's finding of sound as an entity. Similarly, Wittman et al. (2003) encountered student problems when they focused purely on the object-like properties of sound waves, rather than other properties relating to their propagation. Boyes and Stannistreet (1991) uncovered an unusual belief – especially in younger students - in a reverse pathway whereby sound travels from the hearer to the source, rather than source to hearer. I have never encountered that particular notion in many years of teaching Physics, so was interested to see if it surfaced in the present study.

Eshach and Schwartz (2006) used the 11 substance schema of Reiner et al. (2000) to investigate whether students would use these properties to explain sound phenomena. They found that Reiner *et al*.'s (2000) properties were not *all* applicable, but that some of them did apply. In particular students described sound as a substance in terms of it being: pushable, frictional, containable or transitional. They noted too that students' pre-concepts of sound lacked internal consistency. They too cautioned teachers to be careful when describing sound as waves because of the confusion caused between longitudinal, transverse and water waves, echoing Linder's findings and my anecdotal experience. Houle and Barnett (2008) in an interesting study using Urban Bird Communication to teach sound topics, found three main student conceptions, that: sound travels in a sinusoidal wave fashion (as mentioned above and by Linder), sound is a material or a substance and sound is a vibration (the "shaking model" of Hrepic et al. (2010))

Caleon and Subramaniam (2010) surveyed a large group (n = 243) of Grade 10 students in Singapore and identified eleven student conceptions of sound, eight of which had featured in at least two previous studies. Four of these student conceptions were found to be very common namely that: there must be air present for sound to travel, sound can be locked up or trapped in an enclosure if the air is also trapped, increasing the density of a medium makes sound travel faster and, conversely, decreasing the density of a medium makes sound travel faster . Huey-Por Chang et al. (2007) found similarly that students attributed sound leaving a container carried by the air which escaped through small holes in the container-wall. In addition students in Caleon and Subramaniam's (2010)-study asserted that sound, unlike light, cannot be refracted and that wave speed is dependent on wave properties such as frequency or amplitude, with a resulting confusion as to whether pitch is associated with amplitude or frequency. Again many of their students imagined sound as propagating like a (transverse) sine-wave. My anecdotal experience from teaching and performing this show for many years is that students do indeed confuse frequency and amplitude (and pitch and volume), so in the present study questions were designed to test this. I was aware though that their study was done with much more educationally advantaged students than those used in the present study.

Calik (2011) compared different conceptual change pedagogies and their effectiveness and found that a *combination* of different methods (e.g. the use of analogies, computer animations and conceptual change text) was most effective in producing conceptual change in sound-related topics. This provided a sound basis for the multimedia nature of this science show (detailed in 3.3 and in appendix a) incorporating many different teaching methods and modes of multimedia into one presentation.

Finally, Hapkiewicz (1992) compiled the following list of Student Conceptions of sound ("Sound Misconceptions") in the Michigan Science Teachers' Association newsletter, based on teaching experience:

1. Sounds can be produced without using any material objects.
2. Hitting an object harder changes the pitch of the sound produced.
3. Human voice sounds are produced by a large number of vocal cords that all produce different sounds.
4. Loudness and pitch of sounds are the same things.
5. You can see and hear a distinct event at the same moment.
6. Sounds can travel through empty space (a vacuum).
7. Sounds cannot travel through liquids and solids.
8. Sounds made by vehicles (like the whistle of a train) change as the vehicles move past the listener because something (like the train engineer) purposely changes the pitch of the sound.
9. In wind instruments, the instrument itself vibrates (not the internal air column).
10. Music is strictly an art form; it has nothing to do with science.
11. Sound waves are transverse waves (like water and light waves).
12. Matter moves along with water waves as the waves move through a body of water.
13. When waves interact with a solid surface, the waves are destroyed.
14. In actual telephones, sounds (rather than electrical impulses) are carried through the wires.
15. Ultrasounds are extremely loud sounds.
16. Megaphones create sounds.
17. Noise pollution is annoying, but it is essentially harmless.

The first task of this study was to perform a comprehensive synthesis, critical analysis and classification of students' difficulties with the phenomenon of sound based on both the above list and other relevant literature. The findings are presented in table 2.2, at the end of this section. This comprehensive synthesis process yielded an exhaustive list of difficulties which permitted the selection of specific difficulties for further study. The initial synthesis was taken mostly from the literature with some input from my teaching experience.

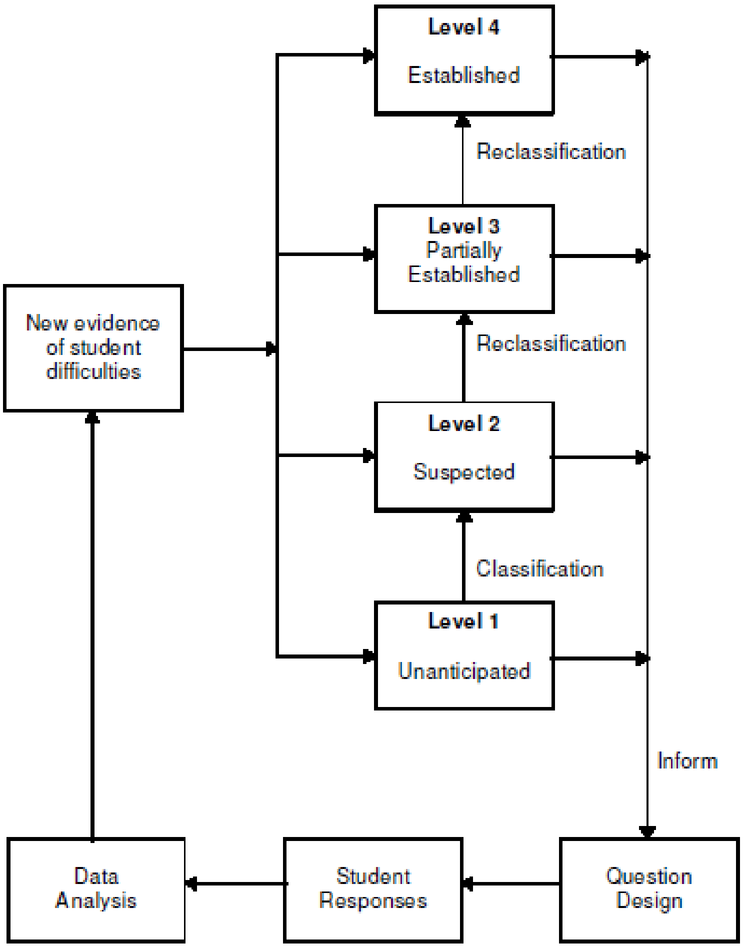


Figure 2.2: Four-level framework for classifying student difficulties, from Grayson, Anderson and Crossley (2001)

The four-level framework of Grayson, Anderson and Crossley (2001) (See Figure 2.2 above) was used to classify the conceptions at different levels, according to the insight one has from research or experience into the nature of the difficulty. Those difficulties firmly established by multiple studies with different groups are classified with a 4 at the top level as established. These were confidently used in devising questions and distracters. Those shown in a single study or in limited studies as only partially established were classified as 3, as only partially established. Those merely suspected from anecdotal information mentioned in single studies (and which I have personally *not* encountered), or from my personal experience (teaching science and doing this show for 20 years) were classified at level 2, as suspected. It is expected that this study will produce some as yet unanticipated conceptions that emerge unexpectedly in the analysis of data, which I will then add at level 1. Finally the difficulties will be reclassified in the light of the results and data, possibly resulting in some of them moving up the list as being more firmly established. A final table of student difficulties (table 5.2) will be given later in Section 5.2 in the light of the results emanating from the data, to answer RQ 2. These processes informed question design and implementation.

In table 2.2 below the far right column lists the references which addressed each student conception. Use the key at the bottom of the page to link to references used above in this section. The 2nd from right column "Hapkiewicz" gives the number used by Hapkiewicz (1992) to refer to that conception – from the list given a few pages back. The other two columns, MCQ and LONG, respectively list the specific multiple choice questions (see appendix b) which address those conceptions and the long, open-ended survey questions (appendix c) also used to address them. The long questions were designed to cover the top 10 student conceptions (and a few others) listed in the table and along with the MCQ to cover all the conceptions presently on the list. There are invariably student conceptions on sound which will not be covered by either test, or during the show, as the very limited scope of this study does not allow it to be exhaustive. These will be of interest in follow up studies. It should also be noted that the concepts related to speed of sound, which give rise to various student conceptions (e.g. 4, 9 and partly 6) will be tested, but deliberately *not* covered during the show, to serve as comparison questions to assist me in linking learning with the show.

Table 2.2: Student Difficulties in Sound from the literature and my experience.

NO	STUDENT DIFFICULTY/ PHYSICS EXPLANATION IN THIS CONTEXT (CONSENSUS MODEL)	LEVEL	Hapkiewicz	Other References *
1	Sound travels as a transverse sine wave	4	11	B, C, E, H, L, R, W
	Sound is a traveling longitudinal-wave which disturbs the particles of the medium.			
2	Sound exists as a microscopic entity either carried by or transferred between molecules	4		E, L, R, W
	Sound is a traveling (longitudinal) disturbance of particles of the medium, not an independent entity			
3	Sound is a macroscopic entity with impetus, in the form of flowing air or a traveling pattern	4		E, H, L, R, W
	Sound is a traveling (longitudinal) disturbance of particles of the medium, not an independent entity			
4	The speed of a sound-wave in a medium is affected by: the force, sound-resistance, separation	4	6, 7	C, L, W
	The speed of a sound wave in a medium depends on the inertial and elastic properties of the medium			
5	Sound can be trapped in a container if the air is trapped (it needs small holes to escape)	4	13	C, P, R
	Sound escapes from a container through vibrations, even if the container is air-tight			
6	Sound can travel through a vacuum and therefore through space	4	6	C, P, R
	Sound requires a medium to propagate and cannot travel through a vacuum			
7	Sound waves cannot be refracted or bent like light	3		C, E
	Sound waves undergo refraction and diffraction in a similar way to light			
8	Confusion between volume and pitch (and amplitude and frequency) and their units	3	4	C
	Volume of a sound (in dB) is associated with the amplitude of the wave, pitch with the frequency (in Hz)			
9	Sound can only travel if air is present (therefore can't travel through liquids or solids)	3	7	C
	Sound needs a medium to travel through but this can be solid, liquid or (any) gas			
10	A human has many different vocal cords to produce different sounds	2	3	
	A human has two vocal cords which vibrate at different frequencies to produce different sounds.			
11	Ultrasounds are very loud sounds	2	15	
	Ultrasounds are very high frequency sounds, which may be loud or soft			
12	Sound travels from the hearer to the source (Reverse sound)	2		B
	Sound waves originate from vibrations at the source and travel to the hearer			
13	Music has low volume (small amplitude) and noise has high volume (large amplitude)	2		My experience
	Music has a clear repetitive wave-form compared with noise			
14	Longer objects vibrate faster, or produce higher notes	2		My experience
	Longer objects vibrate more slowly, producing lower notes			
15	The sound box on a musical instrument is to make the sound clearer	2		My experience
	The sound box on a musical instrument is to make the sound louder (amplify it)			
16	Vibrations and waves are the same thing	2		My experience
	A vibration is a regular movement backwards and forwards			

* Key: B – Boyes and Stanistreet, C – Caleon, E = Eshach, H – Houle, L – Linder (1992), P – Huey Por Chang, R – Hrepic, T – Watt and Russel, W - Wittman

2.7 Design based research (DBR)

In the present study I chose to use DBR as an important guiding methodological framework. Below I describe DBR and present a rationale for why I consider it to be an appropriate framework for the study.

2.7.1 Description of DBR as a research framework

A problem faced in science education is the "credibility gap" (Levin and O'Donnell, 1999), that educational research is often divorced from practice and new research approaches are needed that communicate better with practitioners and lead to the development of "usable knowledge" (Lagemann, 2002). Design based research (DBR, also sometimes known as "design research" or "development research"), which traces its origins back to the early 90's, (Brown, 1992; Collins, 1992) claims to help with this process by mediating both the testing instruments and the intervention itself in an iterative process. The DBR Collective, states that "using methods that link processes of enactment to outcomes has power to generate knowledge that directly applies to educational practice." (DBR Collective, 2003). DBR combines educational research and design of instruction, but performs both simultaneously, instead of sequentially as in empirical research, as the diagram (Fig. 2.3) below illustrates:

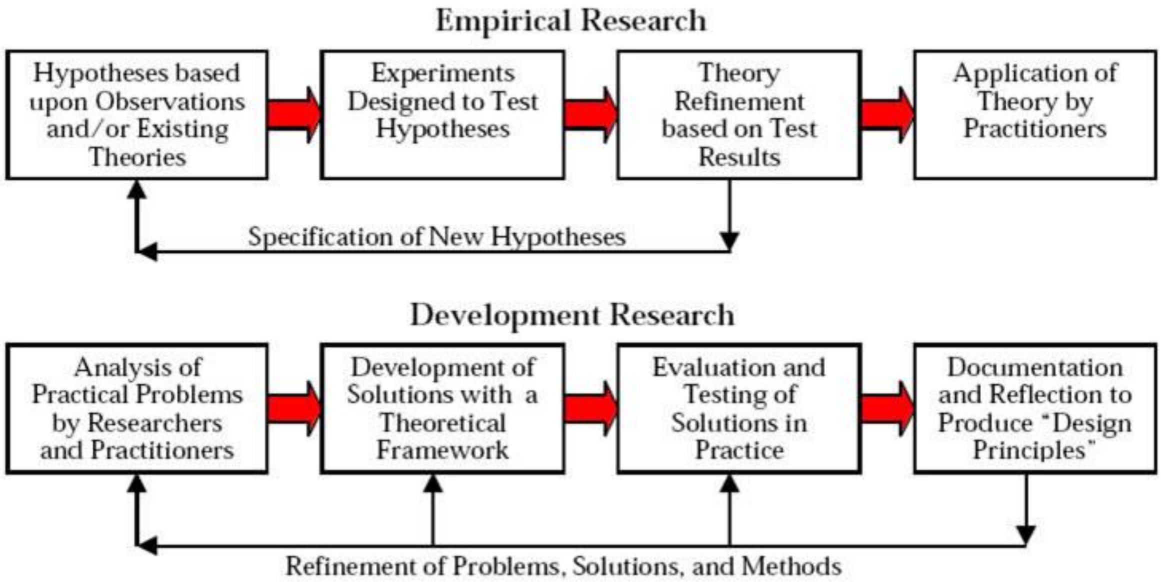


Figure 2.3: Empirical research vs DBR from Reeves (2006)

It is useful to consider the model of Stokes (1997) relating theory and practice in general (Fig. 2.4). When looking at the relationship between technology and basic research in science he suggested a two dimensional model (rather than a linear one of basic research leading to new technology) displaying "quest for understanding and consideration of use" on the two axes.

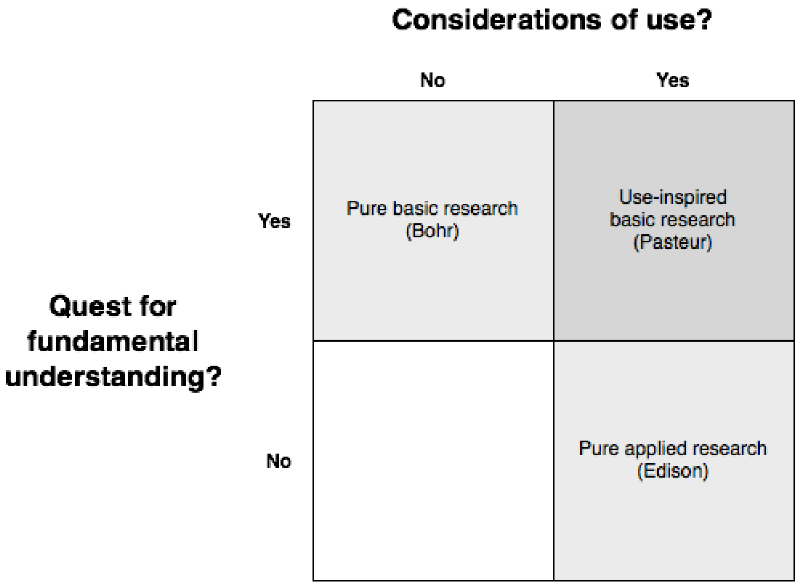


Figure 2.4: Two dimensional model of four quadrants with two dimensions of goals (Stokes 1997).

Prediger et al. (2015) suggest that this model fits the situation in educational research well, and that DBR fits into "Pasteur's Quadrant" as it joins a quest for understanding with considerations of use. This replaces previous curriculum review or instructional theories which were based on a one-dimensional linear model which assumed: "fixed learning goals, ample academic knowledge and directly applicable general theories" (Prediger et al., 2015). These assumptions are not valid for education in South Africa, as was outlined in chapter 1. Finally DBR approaches education from an engineering perspective (Edelson 2002). According to Edelson, DBR focuses on using design processes to create a successful product or useful artifact which can be used practically in instruction (in contrast to the goal of classical research which is new knowledge).

DBR is a useful, practical approach, but has faced criticism from many angles. One of the main areas of criticism is that DBR is not a single, unified research method and that a large variety of research approaches have come to be grouped under the label of DBR. Nevertheless, in an emerging field like DBR, we can comfortably embrace the varieties and look carefully at how different methods are fitted to their contexts. Another major problem is whether it is even possible to pursue a design goal and a theoretical outcome at the same time, with Philipps and Dole (2006) arguing from the perspective of natural science experiments that it is not. Educational research however, especially in the practical setting of the classroom, is a complex process in which variables are more difficult to control than in natural science. DBR emerges as a viable alternative in our quest for better educational tools.

2.7.2 Rationale for using DBR as a framework in the present study

Wang and Hannafin (2005) suggest 5 characteristics of DBR, that it is “pragmatic; grounded; interactive, iterative and flexible; integrative and contextual.” This seemed to me ideal for my needs in this study, as my research is: pragmatic, in that my goal is to solve a real world problem (improving student learning in science shows in science centres). It is grounded, both in literature and the real world context of students in a science centre. It is interactive, in that I span the worlds of researcher and practitioner, ensuring interactive collaboration both by myself and between the networks in which I operate. This study adopts an iterative approach in which findings continually feedback for further improvement to the same show and to the probe. Also, the recursive nature of the process allows for great flexibility and a science centre context is more flexible than the confines of formal education. Furthermore my research is Integrative, in that this study integrates a variety of research methods and approaches, with both qualitative and quantitative probes. Finally it is contextual, in that my research results are “connected with both the design process through which results are generated and the setting where the research is conducted” (the science centre) (Wang and Hannafin, 2005, pg. 11). The four-stage process below, figure 2.5 (from Trujillo, Anderson and Pelaez, 2016) serves as a useful guideline to this process.

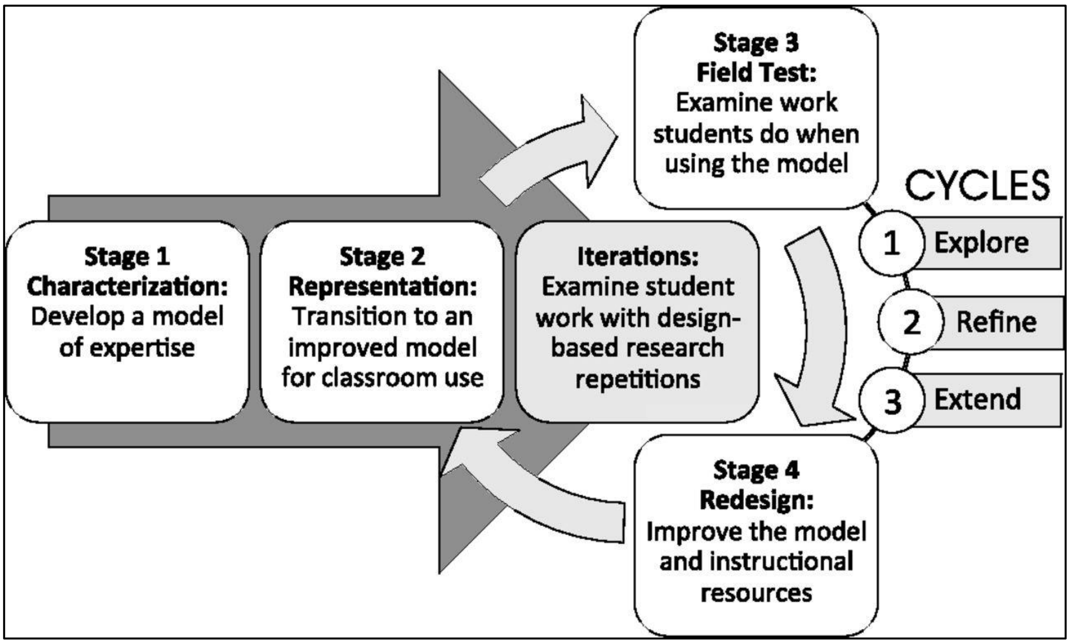


Figure 2.5: Four stage process from Trujillo, Anderson and Pelaez (2016)

It has been used as a framework for this study as follows below (subsequently represented in table 2.3 for ease of use):

1. Stage 1: Characterization: Starting from the theory by performing a literature survey on student difficulties in sound and waves and related fields to develop a model of expertise (Chapter 2), taking into account the contexts influencing the study (section 3.1)
2. Stage 2: Representation: Design of an effective product, artifact or model – in this case a science show ("Good Vibrations" – section 3.2). Design of effective probes to evaluate student use of the model (sections 3.3, 3.4 and 3.5). Pilot testing of the probes and the show (sections 3.4.1 and 3.4.2) (addressing RQ 1)
3. Stage 3: Field test. Validation of suitability of probes to test student understanding of the concepts in the show using student data gathered in the field test (chapter 4). Presentation ("Field Test") of the show and data collection and evaluation (chapter 5). (addressing RQ 2)
4. Stage 4: Redesign: Looking to improve both the show (chapter 6 – addressing RQ 3) and the probe (chapter 7 – addressing RQ 4) in the light of data gathered during stage 3. This will lead to a reiteration of stage 2 (improving the "classroom model") and stage 3 – revalidating and retesting (chapters 8 and 9). These iterations can of course be carried out more than once, and one final iteration is performed with the translation of the show into isiZulu (chapter 10)

Finally, I shall make practical recommendations of how these iterations can ultimately assist science centre practitioners in designing, testing and improving science shows (chapter 11).

To make the terms I use clear: in this study I follow through DBR stages 1 and 2 to develop the *original* show and probes and stage 3 to "field test" them. Stage 4 (redesign) then leads me to produce a *subsequent* show and *subsequent* set of probes for reevaluation. I then perform a second iteration of stage 2 (implementing the new *subsequent* show as an improved model) and stage 3 (field testing it with the new *subsequent* probes to produce new data). A second iteration of stage 4 (redesign) produces a *final* show (in isiZulu). I then perform a third iteration of stage 2 (implementing the new *final* show as an improved model) and stage 3 (field testing it with the same *subsequent* probes as in the second iteration to produce new data). This is where this study ends but the prospect for further improvements through further iterations is discussed. This process is made clearer in table 2.3 below:

Table 2.3: Summary of thesis sections covering each stage of the DBR process (From figure 2.5)

DBR STAGE		ACTIVITY (AND RQ ADDRESSED)	THESIS SECTION	SHOW & PROBES
1	Characterization	Literature survey for theoretical framework Consideration of contexts in development of model	Chapter 2 Section 3.1	ORIGINAL SHOW AND ORIGINAL PROBES
2	Representation	Development of practical model (Science Show) Design of probes for evaluating student response (RQ 1) Pilot test of show and probes	Section 3.2 Sect. 3.3, 3.4 & 3.5 Sect. 3.4.1 & 3.4.2	
3	Field Test	Validation of probe suitability with student data. (RQ 1) "Field Test" of show and data collection & evaluation (RQ 2)	Chapter 4 Chapter 5	
4	Redesign	Redesign of the show in the light of student data (RQ 3) Redesign of the probes in the light of student data (RQ 4)	Chapter 6 Chapter 7	SUBSEQUENT SHOW AND SUBSEQUENT PROBES
2 (2nd)	Representation (2nd iteration)	Transition to the use of an improved model (show) and improved probes	Sect. 3.3.2	
3 (2nd)	Field Test (2nd iteration)	Validation of improved probe suitability with new student data. (RQ 4) "Field Test" of improved show and new data collection and evaluation (RQs 2, 3)	Chapter 8 Chapter 9	
4 (2nd)	Redesign (2nd iteration)	Final redesign (isiZulu language) of the show in the light of new student data (but keeping the same probes) (RQ 3)	Sect. 10.1	FINAL SHOW (ISIZULU) WITH SUBSEQUENT PROBES
2 (3rd)	Representation (3rd iteration)	Transition to the use of a final improved model (*final* show in isiZulu)	Sect. 3.3.2 Sect. 10.2	
3 (3rd)	Field Test (3rd iteration)	"Field Test" of final show. New data collection and evaluation (using *subsequent* probes) (RQ 2)	Sect. 10.2 Sect. 10.3 & 10.4	
4 (3rd)	Further Redesign? (3rd iteration)	Further revisions of the show? (RQ 3)	Sect. 11.4 Future study	?

Chapter 3: Context and Methodology

3.1 Contexts of the study (Stage 1 in figure 2.5)

Having performed a comprehensive literature survey covering all the areas relevant to this study (Chapter 2), it is now necessary to consider the *contexts* of the study, in order to complete the "characterization" stage.

3.1.1 The Unizulu Science Centre (USC) Context:

Falk and Dierking (1992) stressed the importance of the *physical* context of the museum (or Science Centre) in which students learn. Hence it is important to consider Unizulu Science Centre in terms of where it came from, where it is now, what its beliefs are and how it operates.:

1) Early History:
The University of Zululand's Physics Department opened the "Unizulu Science Centre" 30 years ago in November 1986 in a disused laboratory in the Zoology building. The Centre has been open continuously since then and was visited by about 3000 students annually whilst at the University. It comprised some 130 interactive exhibits. I joined the Centre in October 1992 and from then to mid-1997, the Centre was run by the Physics Department and sponsored by the Foundation for Research Development (now the NRF). Since late 1997, the Centre has been run by the University with inputs from various other donors. The Science Centre, with all its exhibits, was moved to its new facility, a disused industrial building in Alton, and opened on the 27th of February 1997. The new facility has allowed for expansions in the programme, visitors and staff, now being visited by over 30 000 students annually.

There have been three expansions to this facility since then. In November 2004, a new 600 square metre building was opened, funded by the Department of Science and Technology. In November 2006, a further extension comprising a 200-seater auditorium (used to present the "Good Vibrations" science show in this study) and a TRAC-Lab was opened, with funding from BHP Billiton (now South 32). In September 2011, a further 600 square metre building was added as an early childhood development centre – the first dedicated children's museum in Africa. The Centre now has over 200 interactive exhibits in 7 exhibit areas covering most areas of Physics (e.g. light, sound, electricity, waves, motion, heat etc.) Present staff consists of a Director (myself), an Operations Manager, and five Team Leaders, in addition to various other facilitators. Since 2008, the Centre has enjoyed the assistance of 6 interns placed and funded by the Department of Science and Technology. In this project, I was assisted by two of the team leaders who handled the surveys while I presented the shows.

2) Location:

Unizulu Science Centre is aimed at Science teachers and students in KwaZulu-Natal schools surrounding the University of Zululand. The focal area stretches roughly from the Tugela River in the South to Kosi Bay in the North, and inland as far as Nongoma. It includes three of KZN's twelve education regions which contain about 550 secondary and over 1000 primary schools. The vast majority of these schools (535) previously fell under the previous Department of Education and Culture (KwaZulu) with the remaining 15 (3 %) being former model C or private schools. Education departments in KwaZulu-Natal were amalgamated in 1995, but historical inequalities in terms of facilities and teachers have not changed much. This is further detailed in Chapter 1 and 3.1.2.

3) Situation:

Education in Northern KwaZulu-Natal generally is in a critical condition, and Science education is no exception, with under-qualified teachers, little or no equipment and many schools not even electrified or supplied with water. The region has also received very little outside assistance in the past, being fairly distant from any large centre. The educational regions covered have been identified in many surveys and studies as the most disadvantaged in the province (and amongst the most disadvantaged nationally.)

4) Goals of the Unizulu Centre:

As early as 1994, the Reconstruction and Development Programme (RDP) confirmed that:

> "Black education, in particular, suffered severe deficits in the areas of Science, Mathematics, Technology, Arts and Culture"

and that we shall have to

> "... take special measures to increase the supply and competence of Maths, Science and Art teachers for schools."

RDP document, sections 3.3.11.9 and 3.3.14.3)

This has been further borne out in numerous other reports (like TIMSS, World Economic Forum, ANA etc.). Unizulu Science Centre has two goals in addressing these critical problems: stimulating an interest in Science amongst school students, and assisting students (especially matriculants) and teachers with the Science syllabus and with examinations. In so doing, it is hoped that this will have the effect of channeling Science students into technological careers.

5) Activities: to achieve these aims, the following activities/programmes take place:
The Science Centre (Exploratorium) forms the base for all activities with more than 200 interactive experiments on display for visiting school-students, in addition to a 200 seater, state-of-the-art auditorium. Most of the students in our contact area have no access to scientific equipment, and gain at best a theoretical knowledge of the subject. The Science Centre seeks to provide a “hands-on” experience of Science for these students. The Science Centre is visited by over 30 000 students each year, mostly from severely educationally and socio-economically disadvantaged, rural schools, who enjoy a four-hour programme comprising a science show (like the one focused on in this study), a small-group workshop, access to interactive exhibits and a visit to the adjacent Careers Centre. In addition the Centre boasts a portable Starlab (planetarium) in which 60 students can enjoy a perfectly accurate depiction of the night skies. Recognising the importance of motivation and interest, especially in teenage students, competitions are run for students to generate enthusiasm for the subject.

There is also a strong outreach component, and quarterly workshops are held to demonstrate practical work to Matriculants. In schools, practical work is examined, but seldom performed. Each Workshop is presented to over 12 000 matriculants at a variety of rural locations. In addition, the Unizulu Science Centre has pioneered the Australian “Science on the Move” programme in South Africa, setting up mobile exhibits in rural schools and taking shows and teacher workshops out of the Centre. In 2007 this programme was augmented by the addition of a “Discovery Bus”:- a mobile unit, containing 40 exhibits.

Teachers are key to any school intervention, so many three-day and one-day courses are run for Science teachers annually to consolidate and broaden their skills, in conjunction with the Centre for the Advancement of Science and Mathematics Education (CASME). A resource centre of equipment, audiovisual aids and reading matter has been set up in the new Richards Bay Centre (run by CASME) to allow teachers access to these vital teaching-aids. Each year 40 Science teachers are taken through the largest Richards Bay industries during a week of their July Holidays, to familiarise teachers with the technology used by the industries, as applications of the school syllabus. For the last three years (2013 – 2015) the Department of Education contracted the Science Centre to train all the mathematics and science teachers in our three local districts, and all the mathematics and science advisors in the province.

Unizulu science centre has always tried to address the real needs of its visitors: Unizulu science centre is situated close to the *world* epicenter for HIV/Aids incidence and prevalence (Hlabisa). Students in local schools are at great risk and Unizulu Science Centre has responded by implementing a comprehensive HIV/Aids education programme, incorporating: multimedia science show, interactive demonstrations, games and exhibits. This science show was the subject of a doctoral research project at the Australian National University. Pre-school education is very problematic in the area, so Unizulu Science Centre has responded by opening the first children's museum in Africa, and enabling Early Childhood Development (ECD) training for local pre-school teachers, in conjunction with TREE (Training Resources for Early Education).

6) Stakeholders:

All aspects of the programme are organized *with* local Science teachers to ensure that the *relevant* needs are met. There is a high degree of cooperation with the local Department of Education offices, especially with the Science Advisors. The National Department of Science and Technology has recognized the work of the Unizulu Science Centre, and made some funding available for expansion and for various DST activities. Networking with other groups is actively pursued to make the most of available resources. The University's Physics Department, who initiated the project, also plays an active and supportive role. Local industries have been extremely supportive both financially and in using their expertise to assist us and to enrich teachers, and new partners are being sought all the time. A number of international colleagues spend time at the Centre each year to share expertise.

7) Recognition:

Unizulu Science Centre has been acclaimed locally, nationally and internationally, winning numerous awards in the field of Science Education. The show in this study, "Good Vibrations" has been presented in Science Centres all over the country for over 15 years and at various national Science Festivals. It has also been presented (in parts) in the USA, Switzerland, Germany, the UK and Austria. The National Department of Science and Technology asked me to present it at the National Launch of National Science Week in 2011, and to represent South Africa by presenting it at the Beijing Science Festival in China in 2012. In addition it has recently been performed at the Abu Dhabi Science Festival in the UAE, and the Science Picnic in Warsaw, Poland. The importance of this from the perspective of the present project is that I have had extensive experience presenting this show and ample opportunity to optimize its activities based on local and international feedback before the performance of the session in this project leading to data collection.

3.1.2 The Student and School Context

Falk and Dierking (1992) also stressed the importance of the *personal and social* contexts of the students visiting the museum (or Science Centre) as major factors affecting their experience. In the light of this, this study focuses on comparing and contrasting three distinct groups of students who attend my Science Centre regularly.

As mentioned in the introduction, South Africa's Education System has been greatly affected by the policy of Apartheid, whereby disparate amounts were spent on different population groups for education. There were multiple departments of Education and differences in spending, standards and some practices. Most white students attended the so-called "model-C" government schools, whereas Black students attended schools under the Department of Education and Culture which were located in Black areas: either in townships or in rural areas. When Apartheid ended with the advent of democracy in 1994, South Africa moved gradually to a single Department of Education representing all racial groups, but historical inequities remained in terms of facilities, class-sizes and quality and availability of teachers. After twenty-five years' of experience at Unizulu Science Centre, I have observed that there are three distinct groups of students who visit my Centre. While it is difficult to fully define these categories, schools can be generally classified as coming from urban, township or rural areas. In the present study I hypothesized that I would find interesting differences and similarities if students were divided in this manner that would allow me to improve our offerings to visitors to the Centre and best address student difficulties. It must be stressed, though, that this decision was based on a generalisation, and that it was important to heed the warning of Eisenhart (2001) that one can run the risk of reinforcing stereotypes based on group membership, and that the knowledge of group patterns should serve as a guideline rather than as a prescription. Nevertheless, in seeking best-practice for Unizulu Science Centre, Lee (2003) stresses that as teachers we require a three-fold knowledge of the science discipline (in this case the physics of sound and waves), the students' language (isiZulu) and their culture (Zulu culture). In the present study, I considered it essential to take cognizance of all three type of knowledge when performing the project.

Urban Schools

Urban schools are typically former model-C schools, which have become quite racially integrated since 1994. They are very well equipped compared with the other schools in the study, and have electricity, water and good facilities generally. Classes in these schools are usually small (25 – 35 students) and science teachers are generally well-qualified with degrees in science. Despite this advantage, it is a sad reflection that even these schools are mostly teaching science without performing student practicals or even teacher demonstrations. The urban school in this study is situated in a residential area of Richards Bay 3 km from the Unizulu Science Centre, and has excellent buildings and grounds. It was built in 1986, so is 30 years old this year. It is attended by 1033 students coming from Black, White, Coloured and Indian population groups and the school is from Grade 8 – 12. The school has 31 classes in total for all grades so the average number of students per class is 33. In the 2015 Matric Exams (just before the final study was performed) this school had 105 candidates for Physical Science of whom 68 passed (65 % pass-rate.)

Rural schools

Rural schools are a relic of Apartheid's "Bantu Education" policies and have a sad history of neglect and under-funding. Sadly, not much has changed since 1994 and schools still battle with inadequate facilities. Most rural schools are ill-equipped compared with the other schools in the study, only some have electricity, only some have water and almost none have science labs or equipment. Classes in these schools are usually large (45 and above students) and science teachers are generally ill-qualified or even completely unqualified in science. These schools are mostly teaching science without performing student practicals or even teacher demonstrations as they lack the basic materials and facilities to do so. The rural school in this study is situated in a rural area just outside Empangeni (less than 10 kms from the town and 25 km from the Unizulu Centre), and has reasonable buildings and electricity. Water is provided by a pump though and there are only pit-latrines. It was built in 1989, so is 27 years old. It is attended by 776 students; all of them isiZulu speakers and the school is from Grade 8 – 12. There are 3 classes per grade so the average number of students per class is 52. In the 2015 Matric Exams (just before the final study was performed) this school had 33 candidates for Physical Science of whom 10 passed (30 % pass-rate.)

Township schools

Township schools fall somewhere between the two extremes outlined above. While they also fell under "Bantu Education", they have generally seen more development since 1994, and are better equipped and resourced, but not as well as urban schools. In addition, qualified teachers are more prepared to live in townships, which are usually peri-urban, than in remote rural areas. Most township schools are moderately equipped compared with the other schools in the study, but most have electricity and water and some have science labs but often with minimal or no equipment. Classes in these schools are usually medium-sized (35 - 45 students) and science teachers are generally qualified in science but usually with diplomas rather than degrees. These schools are also mostly teaching science without performing student practicals or even teacher demonstrations. The township school in this study is situated in Esikhawini Township about 15 kms from Richards Bay and the Centre. It was recently built, in 1998, so is just 18 years old, and has excellent, modern face-brick buildings and good facilities. It is attended by 1184 students, all of them isiZulu speakers and the school offers Grade 8 – 12 since 2014. There are 22 classes across these five grades so the average class size is 54. In the 2015 Matric Exams (just before the final study was performed) this school had 104 candidates for Physical Science of whom 59 passed (57 % pass-rate). The comparative data for these schools is summarised in table 3.1 below:

Table 3.1: Summary of schools' data.

SCHOOLS	**URBAN**	TOWNSHIP	RURAL
GENERAL DESCRIPTION			
Type	Former Model-C	New school	Former Department of Education and Culture
Situation	Suburb of Richards Bay	Esikhawini Township	Rural area outside Empangeni
Distance from SC (km)	3	15	25
Age (years)	30	18	27
Facilities	Very good	Excellent	Poor
Science Equipment	Moderate	Very little	Almost none
Number of pupils	1033	1184	776
Average class size	33	54	52
2015 Matric science candidates	105	104	33
2015 Matric science passes	68	59	10
2015 Matric science pass rate	65 %	57 %	30 %

Jansen (2008), in reviewing Pam Christie's book: "Opening the doors of learning", states that we know from research and observation three reasons that students in township schools under-perform:

- the *time* problem: township students receive a fraction of the teaching time they should
- the *teacher* problem: teachers in township schools lack the depth of subject matter to meet the knowledge demands of the classroom
- the *textbook* problem: schools lack the material resources necessary for effective teaching and learning

These deficiencies which are true in township schools are even more acute in rural schools. These differences are also born out in the students who I see daily in the Science Centre. I have personally interacted with over 500 000 students through the Centre and in outreach and these are my generalised observations:

Urban students usually come from reasonably wealthy homes where they have access to computers, television and reading materials. As a result they are technologically literate and excited by the interactive exhibits in my Centre. Students have good English language skills and a fairly broad general knowledge. They tend to favour individual over group learning which is especially evident on the display floor, where they tend to look at exhibits on their own or in pairs. This tendency also makes them more self-confident and more prepared to ask and answer questions during or after a show. They usually have high expectations of a visit to Unizulu Science Centre and are sometimes quite difficult to handle and raucous.

In contrast, rural students usually come from extremely poor homes without electricity, where they have no access to computers, television and reading materials. They are often quite technologically illiterate although many have cell phones these days. In many ways there is a disconnect between the world they live in and the world of modern science, and between the "primary discourse in the home and community and secondary discourse in school" (Cobern 1998). They are somewhat nervous and fearful of the interactive exhibits in my Centre and need a lot of encouragement to explore these on their own. Students from the rural group also have very weak English language skills as they seldom speak it outside of the classroom. Lack of resources and uneducated parents (or no parents) results in a very limited general knowledge. They tend to favour group learning which is especially evident on the display floor, where they tend to gather around exhibits in groups – especially if there is a staff member present to explain the exhibit. They are very lacking in self-confidence and almost never prepared to ask or answer questions during or after a show. They face the added challenge of coming from a culture which often stresses respect for authority figures (like teachers) over reasoning and that does not encourage enquiry and questioning (Lee 2003). They usually have low expectations of a visit to Unizulu Science Centre and are consequently very easy to handle, being very respectful and grateful.

Township students are harder to classify, but generally fall in the middle of these two extremes. Some of these students can be very similar to urban students especially if they attend a well-resourced school not far from a big city. Alternately, some township schools bear more resemblance to rural schools, being badly resourced and often over-subscribed with too many students for their facilities. Even the best township schools still struggle to compete with ex Model-C schools, and matric pass-rates still reflect the disparities between our schools many years after South Africa obtained democracy.

In the present study it is important not to focus merely on racial groups. Most of the students who attended (107 out of 117, or 91% in the original group, and 101 out of 126 or 80% of the subsequent group) were Black students, but the context of their schools was vastly different. Although these students may share cultural values and language (isiZulu) regardless of socio-economic status, nevertheless socio-economic differences are apparent within the same culture and have great bearing on learning (Lee 1999). Research has often focused more on culture and language than socio-economic status as highlighted by Rodriguez (1998). In addition, as also reported elsewhere in the literature (Rodriguez 1997), little consideration has been given to the prior knowledge and experience of Zulu students in South Africa which makes the "border-crossing" process so much more difficult for them (Aikenhead 1996). This is the challenge for my Science Centre, to assist students to cross borders from their culture and linguistic environments into the culture of schools science with appropriate programmes and materials. Apart from the lowest average score in each of the TIMSS studies mentioned in section 1, South Africa also had by far the greatest spread of scores, ranging from excellent to very poor. This is an indication that South Africa has huge disparities not only in wealth and possessions, but also in the area of educational provision.

At Unizulu Science Centre, we are privileged to receive visits from all three of these school groups. For over 25 years I have noticed distinct differences between the way these different school groups respond to a visit to the Unizulu Science Centre' and I would like to examine this more closely, as it has become painfully clear that one size does *not* fit all- hence the context of the present study is these three groups of students. Activities tailor-made for one group usually do not work well with a different group. This will not just inform, but hopefully also serve to instruct on best practice for each group and in each of the three activities which make up a visit, i.e. shows, workshops and exhibits. Even if other Centres do not have these three distinct groups attending, there are always great variations in student ability and attitude even within one group, and much of my discussion will apply. In his book on the subject, Falk (2009), divides museum visitors into five key types of visitor and describes the unique processes which cause them to visit and revisit, as a basis for making museum visits meaningful for each type. I believe we need to perform a similar separation in South Africa to make our Centres work for all visitors. This belief was investigated in the present study.

3.1.3 The Science Syllabus Context

There have been many changes of syllabus in South African science since democracy in 1994. In 1997 outcomes-based education (OBE) was introduced and lasted until a curriculum review in 2000. This led to the revised National Curriculum Statement (NCS) for Grades R – 9 and Grades 10 – 12 in 2002. From 2012 the two national curriculum statements were combined into a single document known simply as the National Curriculum Statement Grades R – 12 (CAPS, or Curriculum and Assessment Policy Statement). The students in my original study (Grade 9 in 2011) would (mostly) have started school in Grade 1 in 2003, under the NCS syllabus. The students in my subsequent study (Grade 9 in 2016) were in Grade 4 in 2011 and adopted the CAPS curriculum only from Grade 5 onwards.

The previous NCS Syllabus document (see appendix l) had only a fairly vague reference to Sound under the "Energy and Change" section in the Intermediate Phase (Grade 4 - 6), as follows:

> "Sound transfers energy from a vibrating body to our ears. Vibrations travel through a medium, which may be a solid, a liquid or a gas. We hear a change in the rate of vibration as a change in pitch."

There was no reference to any topic related to sound in the Foundation Phase (Grades R – 3) or the Senior Phase (Grades 7 – 9). Thus students may have studied the principles briefly outlined above, but it would have been at least three years prior to my original study, and they would probably not remember them well after that time. In addition, it is possible that the section on sound may have been omitted or badly taught. Both my original and my subsequent study group were taught from this NCS Grade 4 curriculum.

Grade 9 students were chosen for my study especially since the new CAPS syllabus covers sound quite extensively in Grades 10 – 12, and I wanted to test students before such instruction so that I could more easily isolate the effects of attending my show. In summary, the following topics are covered and are listed for completeness:

- Grade 10: Transverse Pulses and Transverse Waves (without direct application to sound)
- Grade 10: Longitudinal Waves, Sound and Physics of Music (directly applicable to this study)
- Grade 12: The Doppler Effect for sound and light

The CAPS document includes the following detail on the Grade 10 sound topics (www.education.gov.za). (Student difficulties from table 2.2 have been added for reference)

Students should be able to:

- Explain that sound waves are created by vibrations in a medium in the direction of propagation. The vibrations cause a regular variation in pressure in the medium (d16)
- Describe a sound wave as a longitudinal wave (d1)
- Explain the relationship between wave speed and the properties of the medium in which the wave travels (gas, liquid or solid) (d 5, d6, d9)
- Understand that sound waves undergo reflection.(d7)
- Understand what are echoes
- Use the equation for wave speed, to solve problems involving sound waves that also include echoes, sonar and bats
- Relate the pitch of a sound to the frequency of a soundwave (d8)
- Relate the loudness of a sound to both the amplitude of a sound wave and the sensitivity of the human ear (d8)
- Describe sound with frequencies higher than 20kHz as ultrasound, up to about 100 kHz (d11)

In addition to the sections above (Grade 10 – 12), the CAPS Curriculum now includes a fairly extensive section on Sound and Waves at Grade 4 level under the "Energy and Change" section of the Natural Science syllabus. While this did not affect any of the students in my study (all of whom finished Grade 4 before CAPS was implemented in 2012), it is a useful consideration for future presentations of "Good Vibrations" – especially to primary school audiences. Topics in this section taken from www.education.gov.za) include the following: (again with student difficulties from table 2.2 added for reference)

Vibrations and sound

- musical instruments make sounds through vibrations (d16)
- the sound always moves outwards from the part that is vibrating (d1)
- we can feel or hear vibrations (d16)
- vibrations travel through materials such as air, water, plastic, metal and wood (d6, 9)

Making sounds

- sounds can be made loud or soft (volume) (d8)
- sounds can be made high or low (pitch) (d8)

Noise pollution

- sound that is loud, unpleasant or harmful to our ears and continues for a long time, is described as noise pollution (d13)
- noise pollution can cause permanent damage to hearing (hearing aids can help people who are hearing-impaired)

Movement and musical instruments (d14, 15)

- many musical instruments (systems) use movement input energy (such as blowing, beating and plucking) to make them work
- many instruments have parts that can move or vibrate
- musical instruments produce sound as the main output energy

This now allows for useful follow-up studies relating the "Good Vibrations" show more closely to the syllabus.

Thus: in addition to considering the students and their schools, I considered the context of the science *syllabus*. For this study I deliberately chose Grade 9 students who had not yet studied sound and waves in class. This would happen in Grades 10 – 12, with most of the relevant work taking place in Grade 11.

3.1.4 The Researcher Context

Whilst one tries to be as objective as possible in any research, one cannot succeed completely and to an extent the researcher will always influence the outcome of any study. Unlike any of the 3 target groups of students, my educational background was privileged with superb science and Maths teachers who not only taught accurately, but challenged and inspired us. I enjoyed many classroom demonstrations of phenomena, and participated in many hands on practicals myself. In all my physics classes at university I enjoyed brilliant lecturers and tutors while lectures were replete with many demonstrations and practical work, which constantly both stretched and inspired me. I subsequently got to teach Physical Science and Mathematics up until Matric level at a good, "all-boys", government high school in East London for three years, before returning to the Science Centre. These experiences in my view amply prepared me for establishing the centre and developing physics exhibits and shows but also served to highlight to me the massive gap between privileged education and the plight of the students visiting my Unizulu Centre. This directly impacted my design of the centre and the science show in this study. Lee (2003) cites extensive literature which shows clear discontinuities between the communication strategies of teachers and their students from diverse backgrounds. He stresses from this literature the importance of "cultural congruence" as a means to promote students' participation and engagement. A struggle I always face in the science centre is getting rural students to participate and to ask and answer questions,

I am personally interested in the "Good Vibrations" Show for many reasons. I designed and developed the show and have been performing it for about 15 years to many thousands of students. As a physicist and piano teacher I enjoy the elements of music and science which interact in the show. It has always been popular with students and the universal appreciation of music (albeit in different forms) allows a good entry point for teaching science. Furthermore the show requires little prior knowledge from students and can be presented (in slightly different forms) to many different age-groups. I have personally presented it to visitors ranging from 3 years old to University Physics Professors. But while I have much anecdotal evidence that the show has been enjoyed and even been effective, it has never been researched. It would be very useful to me, to the whole Science Centre community and to funding agencies to have a better idea of what the impact and effect of this show is with different groups of students. Finally, I have recently had the unique privilege of presenting this show to Chinese, Arabic and Polish audiences (in Beijing, Abu Dhabi and Warsaw) through translators, which has raised many questions about how best to communicate science in different contexts. Despite almost no "cultural congruence" there, I still believe a good science show has the potential to communicate science effectively.

3.2 Description of the science show and rationale behind its design (Stage 2):

Section 2.1 described the three modes of science activity used at Unizulu Science Centre: shows, workshops and interactive exhibits. The same section mentions that science *shows* are the focus of this study and we identify the show then as the external representation (ER) which we shall be investigating. This section then looks at the M-factor – the nature of the *mode* in which sound was presented through the ER (the show: Schönborn and Anderson, 2009).

Science shows at the Unizulu Science Centre are presented by an expert in an auditorium to a large audience of up to 200 children. These shows allow subject matter to be presented in a manner which is dynamic and exciting, as students have relevant experiments demonstrated to them. Falk and Dierking (1992) stressed the importance of the *physical* context of the museum (or Science Centre) in which students learn and the mode of learning is part of that physical context, hence requires careful attention. The shows allow the Science Centre to respond quickly to current topics (like Tsunamis, earthquakes, World Cup Soccer etc.) .Developing a new science show is inexpensive, compared with the very large cost of interactive exhibits, and the multiple sets of equipment needed for workshops.

"Good Vibrations" the show used in the present study, is a dynamic, highly interactive science show, covering the science of sound, and presenting it through the medium of music and musical instruments. It does this in an integrated, multimedia fashion and incorporates the following (Listed in more detail in appendix a):

1. Theme Music
2. Rich Graphics and photos
3. Musical sounds from 30 instruments
4. Volunteer Challenge Demos
5. Whole audience participation Demo
6. Computer based instruments (oscilloscope)
7. Computer simulations
8. Demonstration and playing of 11 normal musical instruments
9. Demonstration and playing of 11 highly unusual musical instruments
10. Demonstration and playing of 3 simple musical instruments which can be made at home
11. Video of unusual and spectacular musical performance
12. Practical applications

The show usually lasts 40 minutes to an hour and has been presented from pre-school to University students, as well as to adult and family groups. The show features a single presenter who presents the dialogues, plays the instruments, performs the demonstrations and operates the computer: showing PowerPoint slides, videos and playing musical sounds. The show thus contains multiple examples of all four of the representational modes mentioned by Gilbert et al. (1998): material (musical instruments, slinky springs etc.), visual (PowerPoint slides, diagrams), verbal (presentation and explanations from the presenter) and symbolic (graphs in oscilloscope form, equations).

As the topic of sound and waves is extensive, the show focuses on just *five key concepts* (KC) which it attempts to communicate in a fun but understandable way. These concepts relate specifically to music and musical instruments which is the medium used to teach sound in this show. In my experience five concepts are about the maximum students can assimilate in a single session without getting too confused. It is naïve to assume one can cover a few weeks' work in a forty minute show and solve many conceptual difficulties. These concepts were also chosen to tie in with the school science curriculum and have been adapted over the years as we have passed from OBE, through NCS to the current CAPS system (see section 3.1.3). These 5 key concepts then represent the C-factor in the CRM model (Schönborn and Anderson, 2009), an abbreviated list of the concepts and knowledge of importance for an introductory understanding of the topic of sound:

1. Sound travels in *longitudinal waves*, which start with a *vibration*.
2. The height of the wave (*amplitude*) indicates the intensity or volume of the sound. (How loud or soft)
3. The width of the wave (*frequency*) indicates the pitch of the sound. (How high or low it is)
4. Music has a repetitive, recognizable wave shape. Noise does not.
5. "Long is low, short is high." A vibrating object which is longer vibrates more slowly (at lower frequency) giving a lower pitch.

To this list I would add other basic knowledge of associated units, like Hz for frequency and dB for volume or sound intensity.

These key concepts are specifically linked to the different elements of the show in the table in appendix a. There is obviously a vast range of instruments one could use in such a show, and so a selection was chosen to include both local and foreign instruments, as well as some highly unusual ones (frog, nose flute, hosepipe trumpet) for their novelty value. All the instruments are used to reinforce the five key concepts listed above, as detailed in appendix a. As the show is presented primarily in Zululand, (although it has been shown all around the world), many local instruments are used in the knowledge that science learning occurs best when linked to the everyday worlds of the students (Aikenhead 1996) and in a context which is linguistically and culturally meaningful and relevant to them (Lee 2003).

> "Recent efforts to provide culturally congruent science instruction indicate that when linguistic and cultural experiences are used as intellectual resources, students from diverse backgrounds are able to engage in scientific practices and show significant achievement gains"(Lee 2003).

As the name of the show suggests, the concept of vibrations is key to the show (KC 1), and is the basis for the main structure of the show. Instruments are divided up according to what actually vibrates in the instrument: i.e. as vibrating solids, skins, strings, air and electronics. In this way the focus in each instrument is brought back to the vibration as the origin of the sound wave (KC 1). The instruments are then individually explored in terms of how they produce notes of different pitch (or frequency KC 3) and how they adjust their volume (or amplitude KC 2). Electronic instruments are merely mentioned as school age students would not understand the complexities of the circuits which produce musical sounds in keyboards, synthesizers or cellphones. Vibrating air instruments get the most attention as this group can be further subdivided into three more groups (flutes, reeds and lip-vibrated) based on how the vibration is produced (KC 1). All instruments are used to reinforce the notion that "Long is low, short is high" (KC 5) and the difference between music and noise is shown using the oscilloscope software (KC 4). This is naturally a very systematic, structured approach which the literature suggests may favour students where it mirrors parenting practices in their homes (Lee 2003) and may thus disadvantage others. With the language problems which are known to exist though, it was decided that having a tightly structured outline would be best for student understanding.

In addition to presenting science in a fun way and helping students to understand sound and waves better, the show could also have other impacts. I would hope that audience members (usually students) would also:

1) See that Science is relevant to all around us and part of all we see and enjoy (like music and musical instruments)
2) Experience that Science is fun and understanding the world is enjoyable
3) Find that Science is surprising!
4) Note that while musical instruments differ, the basic principles are the same.

5) Learn that all cultures have their own musical traditions and instruments- different but not better or worse than others.

6) Be introduced to the Scientific Method: Identify a problem, Form a hypothesis, Devise a method to test that Hypothesis, Gather Results, Conclude

7) Be inspired to go home and try this: Bottle Marimba, Clucking Cup, Straw Oboe

8) Be encouraged to learn to play a musical instrument.

9) Be exposed to a wide variety of instruments.

10) See that our ears and hearing and those of animals are specially adapted to their conditions.

In short, "Good Vibrations" has been enjoyed by thousands of students and other audiences for over 15 years. It has been extensively changed and updated over the years, but never researched properly. There is much anecdotal evidence that the show is popular and that it has made an impact on audiences, but more systematic research is needed to investigate exactly what the effect of this show is for students, especially those from different school groups.

3.3 Research Design and Methodological Framework

3.3.1 Methods: Stage 1 to Stage 3

The research questions stated in Section 1 are reproduced here (for ease of reference). Questions 1 - 2 formed the basis for the *original* study group (first iteration of DBR stages 1 – 3 from figure 2.5).

1) What types of probes would effectively reveal evidence of student learning about sound and any related conceptual and visual difficulties after experiencing a science centre show *on the physics of sound*?

2) How do the students from the *urban, township and rural* school groups compare, after experiencing *the sound show*, in terms of:

a) General attitude;

b) Conceptual and visual difficulties with respect to sound;

c) Prior knowledge and learning relating to the show.

Table 3.2 below summarises the research design and methodology that was used in order to address the specific research questions above. "Good Vibrations" is an existing science show which has been presented in the Science Centre and further afield for over 15 years. The design, testing and final implementation of the probes occurred over a period of a year and a half, from mid-2010 to the end of 2011. The main data gathering took place on 8 November 2011. Data analysis, determination of results, conclusions and implications followed that. More detail follows below table 3.2.

Table 3.2: Schedule of events, instruments and data collected to address each RQ (*original* study):

DATE	ACTIVITY/ DATA COLLECTION PROBE	VENUE	DATA	PARTICI-PANTS	SECTION	RQs.
2000 – 2010 (10 years)	Good Vibrations (GV) show development (Appendix a)	Unizulu Science Centre (USC)	(Anecdotal, visitor surveys)	(20 000 visitors)	3.2	
2010 - present	Literature survey and identification of student difficulties.	USC and elsewhere	Literature, personal experience	Researcher (Derek Fish) and colleagues	2.6	
October 2010	Present GV Show (three times)	Science Unlimited Pieter-maritzburg		(600 visitors)		
October 2010 (after GV show)	**Initial Pilot Study (Appendix g)** (Test for evidence of differences between 3 study groups)	Science Unlimited Pieter-maritzburg	1 page questionnaire: • 6 attitudinal questions • 3 written explanations • 3 drawings	58 students From 3 schools U, T, R Gr 9 - 11	3.4.1 app m	2
March 2011	Present GV Show	USC – Small Theatre		27 students Gr 9 (urban)		
March 2011 (after GV show)	**Open-ended Survey Pilot (Appendix h)** (Trial the probes for the final survey)	USC – Small Theatre	9 long questions • Multiple choice • Written explanations • Drawings	27 students Gr 9 (urban)	3.4.2	1
8th November 2011 (before show)	**Multiple Choice Pre-test (Appendix b)**	USC - Auditorium	13 question MCQ • 3 attitudinal (PRE) • 10 sound concepts (collected with CPS remote clickers)	117 students 3 schools U, T, R Gr 9	3.4.4 4.1 5.1 5.2 5.3	1 2a 2b 2c
8th November 2011	Present GV Show (three times)	USC - Auditorium		117 students Gr 9 U, T, R	3.1.1	
8th November 2011 (just after show)	**Multiple Choice Post-test (Appendix b)**	USC - Auditorium	13 question MCQ • 3 attitudinal (POST) • 10 sound concepts (collected with CPS remote clickers)	117 students Gr 9 U, T, R	app i 3.4.4 4.1 5.1, 2, 3	2a 2b 2c

8th November 2011 (after show)	**Open-ended survey (Appendix c)**	USC – Classroom	9 long questions • Multiple choice (5) • Written explanations(9) • Drawings (8) • Knowledge Source (9)	117 students Gr 9 U, T, R	app j 3.4.3 4.2 5.1, 2, 3	2b
June 2012	**English Language Test: (Appendix d)**	In schools	50 MCQ	60 students (Subset of original study group) From 3 schools U, T, R Now Gr 10	app k 5.1	1 2c
June 2012	Interviews: retention of learning from the show demonstrations	In schools	Informal interviews, notes made on demonstrations remembered.	30 students Now Gr 10	5.1, 4	1 2b 2c

Note: all events occurred in the sequence shown above. For a detailed time schedule for the main data gathering day (8th November 2011), please see appendix f and section 3.4.3.

Regarding the science show to be studied: rather than artificially designing a show for this study, it was felt that it would be more interesting and useful to reflect on an existing show that had been perfected over a long period of time, in this case 15 years (table 3.2). This fits into our DBR framework and reflects the words of Cobb et al. (2003) "The theory must do real work." During this time, although there had never been formal research conducted on this show, it has received much feedback from visitors – especially teachers, and has been continuously changed and improved over a long period. This would make the conclusions and implications more transferable to other shows and other Science Centres. A detailed description of the show and the rationale behind its original design and intentions can be found in section 3.2, and an outline of the show is given in appendix a.

A literature survey was conducted from 2010 to list student difficulties in the area of sound and waves from the literature (see 2.6). This list was augmented with difficulties gleaned from my own experience as a teacher and in presenting this show over the last 15 or more years. From these two sources a table of student difficulties was compiled (table 2.2 in section 2.6). This was used as a basis for the questions in the various probes.

Before starting with detailed probe development and looking for effective probes to use (RQ 1), I felt it necessary to test the *notion* of the study thoroughly first: i.e. will there be discernible *differences* between the three school groups surveyed, which will make the data worthwhile (RQ 2)? I made use of an opportunity to conduct an Initial Pilot Study at the "Science Unlimited" Festival in Pietermaritzburg in October 2010, which is detailed in 3.4.1. The focus of this simple, one-page, questionnaire (appendix b) was to test for *evidence of differences* between the three different groups in the study. Clear evidence emerged (detailed in appendix m), encouraging me to continue. Student answers assisted me in dramatically changing the initial pilot study to make it more suitable.

In March 2011, I trialed the open-ended survey probe with a small group from a single (urban) school at the Science Centre. Here the focus was on the questions themselves (RQ 1), rather than differences between different school groups (RQ 2) and the student answers informed the modification and changes to this Open-ended Survey Pilot – detailed in 3.4.2. The questionnaire used with this group (appendix h) was improved and finalised as the Open-ended Survey (appendix c) to be used at the main data gathering. This open-ended survey attempted to cover the top 10 student difficulties with its questions. One outcome of this pilot was the decision to develop a multiple choice pre and post-test (see 3.4.4) as an addition to the survey to allow for quantitative data as well as qualitative.

On the 8th of November 2011, the main data-gathering took place at Unizulu Science Centre – detailed in section 3.4.3. Students were handled separately in their three groups but all underwent exactly the same programme. They started with a multiple choice pre-test which gathered some attitudinal data (RQ 2a) and then asked 10 general MC questions covering the 5 key concepts in the show and some of the conceptual difficulties (RQ 2b). This MC test (appendix b) was administered with questions on the data projector screen and using the CPS remote clicker system to gather student answers.

After the pre-test (which took only 15 minutes) students watched the science show for 40 minutes. They then immediately wrote a (unanticipated) post-test, with the same 10 general questions, but 3 different attitudinal questions (RQ 2a). The changes, pre- to post-test, were analysed for evidence of student learning during the show (RQ 2c). A detailed, question by question analysis of the MC questions is given in appendix i. Students were then moved from the auditorium to an adjacent classroom where they were seated individually at tables.

In the classroom the open-ended survey was conducted, a 10 page questionnaire gathering biographical data and ethical permission on page 1, and then asking 9 longer questions on concepts in sound (RQ 2 b and c). As detailed in table 3.2, these questions variously requested written explanations, multiple choice answers and drawings. In addition, students were asked to state whether their information in answering the question had come from the show, or was known prior to the show. A detailed, question by question analysis of the Open-ended survey questions is given in appendix j.

Subsequent to the 8th of November 2011, the data from the various probes was synthesised and presented in the results section (5) to answer RQ 2. On the basis of this data, Probes were analysed and validated in section 4 to confirm which probes were effective for analysing a science show (RQ 1). The student difficulties table (table 2.2 in 2.6) was then redrawn to reflect the new information gleaned from this study and answer RQ 2 as (table 5.2 in section 5.2).

After analysing the data, it became apparent that language was a major issue for the students, so a simple language test (appendix d - obtained online from: http://www.transparent.com/learn-english/proficiency-test.html) was administered in June 2012 to test the different groups' abilities, and add to the information available for RQ 2a. As it was not easy to locate all the same students 6 months later (a few had changed schools and one, sadly, had passed away) the test was administered to a representative sample (about half) of each group. The occasion was also used to gather some anecdotal data of what memories students had of the show 6 months later to add to RQ 2a. These "interviews" were very informal and involved talking to some individual students as they finished their English language tests (about 10 of the 20 who wrote). Conditions were not ideal as students were rushing to other classes and it was difficult to isolate students from one another in the venues I was given at the schools. Students were simply asked what they remembered from the show which had taken place 6 months before as it was considered beyond the scope of this study to perform an intensive memory test (as done for example by Sadler (2004)). These "memories" were used to contrast the different school groups, rather than to test for retention of knowledge, which would make an interesting topic for follow-up research. At the end of this intensive process, conclusions were drawn concerning whether research questions had been answered or not (section 11.2) and then the practical implications considered for science centres and for education in general in section 11.5. All of the above formed the schedule for the *original* study group in 2011, and represent stages 1 – 3 in the DBR process (from figure 2.5).

3.3.2 Methods: Stage 4 and Stages 2 and 3 (2nd iteration)

Subsequent to the above, and informed by the data from that initial study, I proceeded to stage 4 in the DBR process: Redesign. Efforts were made to try to improve both the show and the probes (RQ 3 and 4, reproduced below) and a *subsequent* study was conducted in 2016 to measure this. Table 3.3 below reflects these activities and the RQs addressed by them.

3) Building on the data obtained from addressing RQ 2, how can the show be redesigned to improve understanding and minimize conceptual and visualization difficulties with sound?
4) Building on the data obtained from addressing RQ 1, how can the probes be modified so as to be more effective and valid tools for revealing evidence of student learning and any related conceptual and visual difficulties after experiencing *the sound show*?

Table 3.3: Schedule of events, instruments and data collected to address each RQ (*subsequent* study):

DATE	ACTIVITY/ DATA COLLECTION PROBE	VENUE	DATA	PARTICIPANTS	CHAPTER/ SECTION	RQs
2014 and 2015	Analysis of show and probes in the light of student data	Unizulu Science Centre (USC)	Data from original study	Myself, some colleagues and supervisors	6 7	3 4
December 2015	Probe discussion and evaluation: interns	USC	Interns' answer sheets; Recordings transcribed	7 student interns	7	4
December 2015	Pilot test on show improvement	USC	Students' answer sheets	24 Grade 11 pupils (U)	6.3	3
May and June 2016	Presentation of subsequent show and subsequent MCQ and OES testing	USC	13 MCQ pre and post 9 long questions • Multiple choice (5) • Written explanations(9) • Drawings (9)	*Subsequent* study group: 126 students from 3 schools U, T, R Gr 9	8 9	4 3
October 2016	Interviews	At schools	Recordings transcribed	Selected students	7.4, 9.1.2	3, 4

3.3.3 Methods: Stage 4 (2nd iteration) and Stages 2 and 3 (3rd iteration)

Finally, after reviewing the data from the *subsequent* study, I proceeded to stage 4: Redesign for the second iteration. Final efforts were made to try to improve the show (by translating it into isiZulu) for the rural school, but the same probes were used and a *final* study was conducted in October 2016 to measure this. Table 3.4 below reflects these activities and the RQs addressed by them.

Table 3.4: Schedule of events, instruments and data collected to address each RQ (*final* study):

DATE	ACTIVITY/ DATA COLLECTION PROBE	VENUE	DATA	PARTICIPANTS	SECTION	RQs
July and August 2016	Development of isiZulu Show	USC	New PowerPoint and script	Myself and isiZulu colleagues	10.1	3
October 2016	Presentation of isiZulu Show and testing with subsequent MCQ and OES	USC	13 MCQ pre and post 9 long questions • Multiple choice (5) • Written explanations(9) Drawings (9)	*Final* group: isiZulu Group 42 Grade 9's	10.2 – 10.4	3
October 2016	English Language Test	At rural school	50 MCQ	30 of the 42 students from the isiZulu group	10.3	4

3.4 Data collection instruments (Stage 2)

3.4.1 Development and Design of Probes: Initial Pilot (October 2010)

Before starting with detailed probe development, I felt it necessary to test the *notion* of the survey thoroughly first: i.e. will there be discernible differences between the three school groups surveyed which will make the data worthwhile? An opportunity to conduct such a survey arose as I was presenting: “Good Vibrations” at the “Science Unlimited” Festival at the Royal Show Grounds, Pietermaritzburg on Tuesday 19th and Wed 20th October 2010. This was a 45 minute presentation to varied audiences comprising primary and high school students and their teachers, as well as the general public. Audiences were about 200 – 250 people per show. While not a strictly controlled environment, it nevertheless offered me the opportunity to conduct some surveys. My supervisor, Prof. Trevor Anderson, was also able to attend the show and give input into the resultant probe analysis and redesign.

After each of three presentations of Good Vibrations, I asked permission (from the teacher) to conduct a survey with a small group of about 20 students from different schools who remained behind after the show. Students filled in the questionnaire (attached as appendix g), taking about 20 minutes to do so. The biographical information for the three groups is given in table 3.5 below:

Table 3.5: Data for schools in Pilot Study

School type	Situation	Grade of students	Gender	Number	Ave. Age
Urban School	Extremely privileged	Grade 9	Boys only	21 students	15.0
Township School	Close to Pietermaritzburg	Grade 11	Boys and girls	19 students	17.1
Rural School	Outside Howick	Grade 10	Boys and girls	18 students	15.9

Conditions for the survey were not ideal, as inevitably some other students hung around in the venue after the show making it difficult to control the students doing the survey. In addition, time was very short and students were rushed. As they were attending a festival, students appeared more concerned about the next show they were due to attend than about filling in the whole questionnaire diligently. Nevertheless this initial pilot study, yielded useful information that informed the design and development of the main study. Detailed results for each question are presented in appendix m.

This pilot study appears to have achieved its main aim of showing clearly different results between the three school groups surveyed. Attitudinal questions (Q 1 - 6) showed interesting variations in statements of enjoyment of the show, novel experiences and intention to pass on information learned. While these questions yielded some interesting insights they do not seem to be as useful in deciding whether a science show has been worthwhile. The questions themselves had too much overlap and would probably be better implemented via a multiple choice instrument. Conceptual questions which were scored on a 6 point scale of ability (Q 8, 9 and 11) showed urban students clearly performing better than township, who were a little better than rural. The scoring system was problematic though, and in future it was felt an inductive analysis would work better. The two questions requiring drawings (Q 7 and 10) yielded fascinating data, and have great potential for further investigation. In addition, as long as the question is clearly stated and understood, the answer is language-free and should help students with poor language ability (like the rural schools) to answer clearly. I would recommend from this that almost every question in the survey include a drawing, as these have yielded very interesting data in this pilot study. There was also a marked difference in writing ability and depth of answers offered. To this extent it has achieved its aim of showing discernible differences between the groups and I am confident that a broader study of this kind will yield fruitful results.

Practical recommendations for final survey design and implementation

1. Ensure plenty of time for survey so students are not rushed
2. Ensure a quiet venue where students will not be interrupted
3. Gather attitudinal data through simple multiple choice questions (administered by CPS)
4. Be aware of severe language problems – especially for rural students.
5. Structure more questions (or even *most*) like q's 7 and 10, requiring written explanation *and* a drawing.
6. Allow for a simple pre- and post-test (probably multiple choice) to compare learning in the show within and between school groups.

3.4.2 Development and Design of Probes: Open-ended Survey Pilot: (Mar 2011)

Having thoroughly tested the *notion* of the survey in the previous section, and finding clearly discernible differences between the three school groups surveyed, I moved on to look in more depth at the open-ended survey which would be administered after the show. I presented Good Vibrations to 29 Grade 9 students (a mixture of boys and girls) from a small school close to my Science Centre in March 2011. I asked permission (from the teacher) to conduct a survey with these students as they remained behind after the show. Students filled in the questionnaire (attached as appendix h) after the show, taking about 30 minutes to do so.

Conditions for the survey were better than previously, as there were no other students present and we were not rushed. No attempt was made to compare different types of school groups here, as this study was simply to test the probes themselves and to see whether the questions were clearly understood, and yielded useful answers. Following on the conclusions of my previous pilot study, nine questions were devised testing concepts in sound and waves and requesting a combination of written explanations, drawings and multiple choice answers. This probe was extensively discussed with both my supervisors, with particular input on the Physics in the questions from Dr. Saalih Allie. Instead of a traditional multiple choice format a cartoon format was used (as in Allie and Buffler 1998) asking students to agree with an opinion expressed by a character illustrated. The characters were deliberately free of gender and race, and the wording kept very short and to the point to minimize language problems.

These results were not analysed in detail, but rather studied to see where questions had been misunderstood and could be improved. An item analysis was performed on the multiple choice elements. Questions included drawings as well as written explanations as language is a big issue for students around Zululand, and drawings are essentially language free (as long as the question is well understood). Kose (2008) confirmed that drawings are a simple research instrument that allows for international comparisons and brings out ideas which a student may struggle to express in words. Questions were chosen based on identified student difficulties in sound and waves (from 2.6) in order to cover at least the top 10 student difficulties, but this will be further elucidated in 3.4.3. A detailed analysis of each question is given as appendix n.

In general this second probe seemed to have achieved its goal of testing conceptual understanding of sound and waves after seeing the show: "Good Vibrations." Notwithstanding the comments made on individual questions in appendix n, most students understood most questions and were able to attempt answers and drawings. It is interesting to consider specifically the results for the answer selection in the first 6 questions (A – D), and the figures produced by an item analysis:

Table 3.6: Item analysis results for pilot open-ended survey questions 1 – 6:

QUESTION	**1**	**2**	**3**	**4**	**5**	**6**	**AVE**
DETAIL	Sound through wall	Loud hurts ears	Men have deeper voices	Guitar has wooden box	Which phase sound fastest	Hear sound in space	
NUMBER RIGHT ex 29	5	12	26	15	12	12	**13.7**
DIFFICULTY INDEX (p)	0.17	0.41	0.90	0.52	0.41	0.41	**0.47**
DISCRIMINATION INDEX (DI)	0.29	0.86	0.29	0.29	0.71	0.43	**0.48**

The two shaded columns (questions 3 and 4) represent the two concepts which were *explicitly* dealt with in the show:

3. Boys' vocal cords become thicker, longer and looser as they become men
4. The sound box of a guitar amplifies sound coming from the vibrations of the strings

These obtained the best scores from the students, and the scores reflected the attention given to the concepts in the show. They also produced low DI values, but this was probably due to most of the students in a small class knowing the correct answer. Men having deeper voices than ladies (Q 3) was mentioned many times and related to a demonstration on the vibrations of guitar strings (with men depicted as "bass guitars" with thicker strings). The sound box of a guitar (Q 4) was demonstrated by placing a small imbira (music box) on top of it and following up with an explanation of a musical wooden jewelry box. Thus it is felt that the questions do reflect what students have gained from the show. With this evidence, it was decided to use the 9 question probe for the open-ended survey, making minor (but important) changes to individual questions which are summarised in table 3.9 in the next section. This final survey would be used as a post-test only, with a brief multiple choice test administered as both a pre- and a post-test for comparison.

3.4.3. Final Design and Implementation of Open-ended Survey

The main original open-ended survey was conducted on the 8th of November 2011, at Unizulu Science Centre, with the following three groups totaling 117 students and three teachers. (I had hoped to get three groups of 40 students but the rural school arrived three short)

Table 3.7: Biographical details of schools in *original* study:

School type	**Situation**	**Grade**	Boys	Girls	Total	**Black students (isiZulu speaking)**	**Other students**	**Ave-rage Age**	**Survey Nos.**
Urban School	Model C in Richards Bay	9	12	28	40	30	10	15.05	101 – 140
Township School	From Esikhawini	9	20	20	40	40	0	15.30	301 - 340
Rural School	Rural area near Empangeni	9	18	19	37	37	0	15.62	201 - 237
		TOT-ALS	**50**	**67**	**117**	**107**	**10**	**15.3**	

The schools were bussed in to the Science Centre to ensure that they all arrived on time and that conditions were identical. (Logistics are outlined in a table as appendix f) They were seen in the order: urban, rural, township and the same colour coding (black for urban, red for township, green for rural) is used throughout in graphs for ease of identification. Every page of the written survey was numbered and coloured as in the right hand column, and the same numbering was applied to the remote clickers (CPS system) for the pre- and post-tests to ensure that results could be easily compared and correlated. The schedule for each group was as follows:

Table 3.8: Schedule for each group in *original* study

No	Event	Format	Venue	Duration	Data
1	PRE-TEST (Appendix b)	13, four-part MCQ on PowerPoint, answered with clicker	Auditorium	15 mins	CPS pre-test excel file
2	GOOD VIBRATIONS SHOW	Science Show	Auditorium	45 mins	(Informal observations of attitude and involvement)
3	POST-TEST (Appendix b)	13, four-part MCQ on PowerPoint, answered with clicker	Auditorium	15 mins	CPS post-test excel file
4	OPEN-ENDED SURVEY (Appendix c)	9 Questions requiring: • Explanations • Drawings (8) • Multiple choice (5) • Source of knowledge	Classroom	30 mins	Biographical Data Questions 1-9 (all on 10 page numbered form)
5	COOLDRINKS AND RELAX		Exhibit Area	15 mins	
			TOTAL	2 hours	

I personally administered the pre-test and post-test and presented the show to all three groups, to ensure consistency. The show was presented from a script (not my usual practice) and the same PowerPoint slides used. The show was presented in English and no allowance was made for language difficulties, and no extra explanations given even if it was seen that students were struggling to keep up. The show was precisely timed at 45 minutes and no extra time given to any group. The open-ended survey was conducted by staff of Unizulu Science Centre (while I started with the next group). It was very closely invigilated to prevent copying and students were encouraged to complete the surveys and not to leave anything out. This was of course not foolproof, and a few questions and drawings towards the end were left out by some students. Conditions generally were as ideal as possible, with all three groups experiencing identical conditions, the same schedule and the same show from the same person. The open-ended survey was written in a comfortable, air-conditioned venue, with sufficient time to avoid too much rushing.

The cover page of the open-ended survey gathered basic biographical data (Gender, School, Name and Age) and was signed by students to indicate full understanding and compliance with the conditions of the survey. Full ethical clearance was obtained from UKZN (appendix e) and agreement from the relevant schools. Students were encouraged to complete the exercise with the promise of a reward at the end – a cool drink. (Abell 1992). Wittman et al. (2003) confirmed that teaching by way of lecture instruction alone had little effect on student learning and understanding of sound. I was interested to see what the response was after a science show which contained many interactive demonstrations.

The open-ended survey (comprising 9 questions) was separated into questions to be comprehensively analysed. Some general comments on the probe itself as reflections after analysis follow. In Section 3.4.2 and appendix n, I discussed how the questions were altered after the pilot survey with 29 local students. The changes made appear to have been very effective as there seemed to be limited misunderstanding of the questions. Apart from changes in the order of the questions, the table below summarises the main changes made:

Table 3.9: Summary of changes made to open-ended survey after pilot study:

NO.	FINAL SURVEY QUESTION	PILOT Q. NO.	EXPLAIN	DRAW	CHOOSE ANSWER	CHANGES TO ORIGINAL
1	How does sound travel through the air?	7	YES	YES		"Bugs" removed
2	How does sound travel through a wall?	1	YES	YES		"Confused" removed No multiple choice Drawing required
5	In which phase does sound travel the fastest?	5	YES	YES	YES	Explanations removed from bubbles. Drawing required
8	Why do men have deeper voices than women?	3	YES		YES	Distracter C changed

Questions seemed to be clearer and better understood after removing the reference to "Bugs" (Classmate's name) and taking away the statement "Explain to a friend *who is confused*" from all the questions. Question 2 definitely worked better without the multiple choice options and with the request for a drawing. Question 5 also benefited from simpler multiple choice options (phase only, no explanation) which then required the students to come up with explanations (which they mostly couldn't). Again the request for a drawing here made demands on the students and very few were able to produce anything satisfactory, showing very little understanding of this phenomenon. In Question 8, the distracter (C) which had had no hits was changed and the new option resulted in 16 choices, so clearly was a more relevant choice for students.

The open-ended survey (attached as appendix c) adopted a three or four tier diagnostic approach similar to that of Caleon and Subramaniam (2010). As they discussed – the simple two-tier approach which is usual has two main drawbacks, in that it cannot differentiate between mistakes due to lack of knowledge and those due to alternative (or mis-) conceptions and that it cannot differentiate between correct responses due to full understanding and those due to guessing. As I was testing Grade 9 students who had not yet studied sound, I decided not to utilize the confidence rating they (Caleon and Subramaniam) used, but rather to try to discern whether students had used knowledge known before the show or gained in the show. This gave me a number of different tools to use in assessing the students, which were detailed in table 3.2. Because of concerns over the English language abilities of students, it was decided to ask for explanations via drawings as well. In many ways drawings remove the language issue (assuming the initial question is understood) and allow me to probe the visual dimension of student knowledge. In addition, they provide a window into students' thoughts and feelings as they reflect an image of their mind. (Tomas 1990) – as quoted in (Kose 2008), who proposed a five level framework for analyzing student drawings: no drawing, non-representational drawings, drawings with misconceptions, partial drawings and comprehensive representation drawings. His study was with third year college students, and I felt this framework was too comprehensive for this study, so used a more simplified form of inductive analysis.

The final 9 questions were carefully chosen to cover the top 10 student conceptions from table 2.2 in section 2.6. These are illustrated in table 3.10 overleaf. Finally, all the data from this survey was assimilated and analysed in the results section: Chapter 5. In terms of the CRM model (Schönborn and Anderson, 2009) these OES questions (administered only as a post-test) are an attempt to test some aspects of the R-factor – the total reasoning ability or skills the student would need to interpret the show (ER). This would include a student's ability to reason with his/her conceptual knowledge of relevance to the ER (the R-C factor) including assimilation, accommodation and integration of new knowledge from the ER, as well as transfer. In addition it would include the R-M factor: the student's ability to reason with the ER itself and its graphical features. This may include skills like observation, visualization, formation of simple mental models, knowledge transfer, real life application and deduction.

Table 3.10 Student difficulties and questions addressing them

NO	DIFFICULTY/ CORRECT PHYSICS	LEVEL	MCQ	OES
1	Sound travels as a transverse sine wave	4	2	1,2,3
	Sound is a traveling longitudinal-wave which disturbs the particles of the medium.			
2	Sound exists as a microscopic entity either carried by or transferred between molecules	4		1 (2,3)
	Sound is a traveling (longitudinal) disturbance of particles of the medium, not an independent entity			
3	Sound is a macroscopic entity with impetus, in the form of flowing air or a traveling pattern	4		1 (2,3)
	Sound is a traveling (longitudinal) disturbance of particles of the medium, not an independent entity			
4	The speed of a sound-wave in a medium is affected by: the force used to create the wave, the "sound-resistance" of the medium, separation of the molecules or compressibility	4	3	5
	The speed of a sound wave in a medium depends on the inertial and elastic properties of the medium			
5	Sound can be trapped in a container if the air is trapped (it needs small holes to escape)	4		2
	Sound escapes from a container through vibrations, even if the container is air-tight			
6	Sound can travel through a vacuum and therefore through space	4	3	9
	Sound requires a medium to propagate and cannot travel through a vacuum			
7	Sound waves cannot be refracted or bent like light	3		3
	Sound waves undergo refraction and diffraction in a similar way to light			
8	Confusion between volume and pitch (and amplitude and frequency) and their units	3	6,7,10	6, 7
	The volume of a sound (in dB) is associated with the amplitude of the wave, the pitch with the frequency (in Hz)		4,5	
9	Sound can only travel if air is present (therefore can't travel through liquids or solids)	3	3	9
	Sound needs a medium to travel through but this can be solid, liquid or (any) gas			
10	A human has many different vocal cords to produce different sounds	2		8
	A human has two vocal cords which vibrate at different frequencies to produce different sounds.			

11	Ultrasounds are very loud sounds	2	10	
	Ultrasounds are very high frequency sounds, which may be loud or soft			
12	Sound travels from the hearer to the source (Reverse sound)	2		1
	Sound waves originate from vibrations at the source and travel to the hearer			
13	Music has low volume (small amplitude) and noise has high volume (large amplitude)	2	8	4
	Music has a clear repetitive wave-form compared with noise			
14	Longer objects vibrate faster, or produce higher notes	2	9	
	Longer objects vibrate more slowly, producing lower notes			
15	The sound box on a musical instrument is to make the sound clearer	2		6
	The sound box on a musical instrument is to make the sound louder (amplify it)			
16	Vibrations and waves are the same thing	2	1	
	A vibration is a regular movement backwards and forwards			

3.4.4 Multiple Choice Pre and Post-Test

After the Initial and Open-ended Survey Pilots, it was felt that useful data could be obtained from a pre- and post-test in multiple choice format. In addition that this would be an effective way to test the C-factor of the CRM model (Schönborn and Anderson, 2009): a student's *prior* knowledge of concepts represented by the ER (and outlined in the 5 key concepts in section 3.2), via a simple pre-test. The same questions, administered as a post-test, would also aid in evaluating the R-C factor, looking at assimilation, accommodation and integration of new knowledge from the ER.

The Pre and post tests were thus presented in PowerPoint form, question by question, on the main screen in the auditorium. The tests are attached as appendix b. Questions 1 – 10 were identical in both pre and post-tests and tested basic concepts in sound and waves. Questions 11 – 13 were different in each test and looked at attitudes and intentions. Students were not told about the post-test beforehand so were presumably not watching the show with the specific aim of looking for answers to the pre-test questions, although the potential for students to seek answers to the pre-test during the show cannot be ruled out.

The CPS (Classroom Student Response) or Clicker system was a great asset in data collection and I would strongly recommend its use to researchers. It allowed instant collection and analysis of data, prevented copying or cheating and ensured that no questions were left unanswered. Furthermore its novelty kept the students interested and alert during tests. (Brenner 2010) reported that students liked the way that clickers force interaction, the way they show them if they have grasped a concept, their anonymity and the instant feedback and reports they can generate. (Mazur 2009) writes that data obtained by him and colleagues worldwide across any different disciplines show that learning gains nearly triple when the approach to Physics teaching focuses on the student and on interactive learning through the use of clickers.

As detailed in section 3.2, the show “Good Vibrations” focused on 5 key concepts in the topic of sound and waves (the C-factor). The 10 pre- and post-test questions were designed to cover these 5 concepts, test some baseline knowledge (like units) and test *some* of the misconceptions from the literature. It was felt, though, that the misconceptions would be better probed through the longer questions and drawings (as shown in table 3.10) on the previous page. Table 3.11 below explains the choice of the multiple choice questions.

Table 3.11: Multiple choice questions with key concepts and student conceptions addressed.

MC Q NO	QUESTION (Appendix b)	KEY CONCEPT (from list above)	STUDENT CONCEPTION (from table 2.2)
1	What is a vibration?	1 b	16
2	How does sound travel (type of wave)?	1 a	1
3	Where does sound travel fastest?	COMPARISON QUESTION	4, 9
4	Unit for volume?	Baseline knowledge	8
5	Unit for frequency?	Baseline knowledge	8
6	Wave property related to volume?	2	8
7	Wave property related to pitch?	3	8
8	How does music's wave differ from noise?	4	13
9	How will a longer string vibrate and sound?	5	14
10	Why can dogs hear sounds humans can't?	Gen. knowledge(3)	

As seen in the table above, questions 1 – 3 were set specifically to investigate common student conceptions of sound (16, 1, 4 and 9) from table 2.2 in section 2.6. Question 3 probed two common student conceptions identified from the literature; especially conception 9, that sound can only travel through air and therefore will not travel through liquids or solids). It was also used as a comparison question, to see if the learning that occurred could be specifically linked to the show, or was just coincidental, or the result of a gain in confidence as time was spent in the science centre. As such it was specifically *not* covered in any way in the show, and no attempt was made to correct any student conceptions on this topic. All other questions were directly covered during the show, so that I could test for improvement in student knowledge and understanding.

Questions 4 and 5 simply questioned some baseline knowledge: whether students knew the units of measurement of volume and frequency, but also assisted me in determining whether students did indeed confuse frequency and amplitude (and pitch and volume) – item 8 on my student difficulty table. Questions 6 and 7 probed whether students learned two of the other important key concepts in sound and waves: i.e. which wave property is related to which physical occurrence (volume and pitch). Questions 8 and 9 probed whether students learned why music differed from noise, and how instruments produced different pitches. Question 10 was a general interest question although indirectly related to key concept 3. Distracters used were chosen from the literature (student conceptions table 2.2) and also from my 20 years' experience with this show and with teaching science.

3.5 Data processing (Stage 2)

As mentioned in the sections above, it was decided to perform a mixed-mode study, as it was felt that the open-ended survey administered as a post-test only (3.4.3) was more suitable for determining student difficulties (R-C and R-M factors), and a multiple-choice pre-test (C-factor) and post-test (3.4.4) more suitable for gauging learning during the show (R-C factor) and some baseline knowledge. Furthermore it was felt that this mixed mode would be more likely to be effective for a large group from various backgrounds, as different students would probably perform better at different types of test. Practically as well, the multiple choice elements allowed for swift data collection and analysis, while the open-ended survey required more time and effort to collect and to analyse, but allowed for more depth in the student answers.

3.5.1 Inductive qualitative analysis

This analysis was performed on the Open-Ended Survey (OES) questions which were free-response in nature, allowing students to speak their minds more freely, and to show their own unique conceptions and understanding. The aim of the inductive analysis was to condense the students' responses into a more manageable format, allowing patterns to be seen amongst the data. Student responses were coded with the same codes used for the same difficulties so that I could get an idea of their frequency. This also assisted in comparing and contrasting across different school groups, where the manner of answering was often quite varied. This analysis sought to reduce the number of categories used to a manageable number – usually fewer than ten. Where necessary, a two-level coding was used to reveal further details of the student answer. For example, in the drawing for question 2: (sound traveling through a wall) a first code was used for the *type* of wave drawn, and a second code for the *way* it interacted with/ passed through the wall. This inductive analysis was used both for the drawings and the written answers required by the open-ended survey questions 1 – 9. The "choose-an-answer" part of open-ended survey questions 5 – 9 was analysed using quantitative statistical means. Through the OES it was expected that students would translate their mental model to an expressed model which could then be compared with the consensus model (expressed in the propositional knowledge statements. (Gilbert et al., 1998)

3.5.2 Quantitative statistical analysis

Where the data allowed, quantitative statistical analysis was used to analyse it. This applied to the "choose-an-answer" part of open-ended survey questions 5 – 9 and to the multiple choice pre and post-tests. Data for the multiple choice sections were gathered using the CPS “clickers” which greatly simplified analysis. It was fed directly (by the system) into an excel spreadsheet from where it could be analysed. This allowed for very simple distracter analysis, and along with the key for each question, a simple calculation of how many students had answered the question correctly. Comparison of pre- and post-tests was also simplified by the system and allowed me to look at individual and group performances.

Validation of the questions was achieved by performing an item analysis on each question. This involved:

- Calculating an item difficulty index of each test item. This is a measure of the proportion of students who answered the item correctly. This would assist in deciding if the questions were too easy or too difficult, and whether there was any improvement pre- to post-test.
- Calculating the item discrimination index of each test item. This is a measure of how well the item is able to discriminate between the top and the bottom students. It is calculated by subtracting the number of correct answers for the bottom 25 % from the number correct for the top 25 % and dividing by the number of students in that group. This assists in weeding out questions which are too hard, too easy or just confusing.
- Doing a distracter analysis for each test item. By considering what proportion of the students chose each distracter, I can see whether our distracters worked well or not.

In addition not just the mean, but also the spread of the results was considered to ensure that there was variation in these results. When considering learning during the show, the pre- and post-test scores were subject to a Student T-test to ensure that the change was statistically significant and not just the result of chance.

3.5.3 Limitations of the methods

Any form of analysis will always have some limitations, so one should consider what these will be here. While the inductive analysis assists greatly in managing the data, the process of coding and classifying can be too rough, and can lead to student differences being grouped together which still have subtle variances. These can then become obscured by the classification process. In addition there is always some degree of subjectivity in this classification process, and with the large amount of data coming in and the limitations of this study, it is difficult to avoid this. Another problem (already evident from the pilot study and survey) is that student drawings, and even more so – written answers, can be so unintelligible that classification becomes impossible. Follow up interviews could be helpful here but would be very time-intensive. In the current circumstances, I shall simply have to allow a category for those drawings or answers which are unintelligible to us, which will effectively be treated like those left blank. In addition the very weak standard of English amongst students places imitations on their understanding of the questions, as well as the quality of their written answers. Also the whole notion of a survey (which to students is like an exam) can be so intimidating that it prevents students from answering freely. Finally, as will be discussed later, the open-ended survey takes a long time to answer making it unsuitable as a pre- and post-test. In addition the marking and coding takes a long time which makes it less practically useful to science centre staff already under time pressure. Finally a number of factors militate against students effectively translating their personal mental models into expressed models which we can access.

While some of the problems and limitations mentioned above do not apply to multiple choice questions, they are limited in that students are not free to express themselves in their answers. Students have to choose from one of four possible answers, and this may well mask interesting student responses and difficulties which could otherwise have been revealed. While the provision of options may help some students to answer, there will still be a limitation for those who are weak readers of English. Nevertheless item analysis allows one to develop and refine good questions which can give valid results.

Chapter 4: Instrument Development and Validation: Original study (Stage 3)

Research question 1, namely: “What types of probes would effectively reveal evidence of student learning about sound and any related conceptual and visual difficulties after experiencing a science centre show *on the physics of sound*?” was addressed in this chapter. Based on the experience gained from the pilot-studies, and given the varied nature of the school groups, it was decided to use a range of both qualitative and quantitative probes to investigate student understanding of sound. Before using student responses to these probes to characterize and contrast the school groups, it was necessary to establish probe validity. Graue (1993) stressed that: “Choosing the appropriate assessment strategy is a validity concern; the tool must be relevant to the task at hand…”. The issue of validity will be discussed in terms of three questions:

1. What *type of knowledge* is the question valid for probing?
2. Why is this question asked? For what *purpose* is it valid? Who will use the data?
3. For whom is this question valid i.e. *what kinds of student*? In my case, is it valid for all three groups, or only for some of them?

In terms of these questions, the probes are summarised in table 4.1 with respect to their *initial intentions.* It was hoped that all tests would prove valid across all three school groups as the same tests needed to be applied to all three so that I could compare and contrast the groups. This table gives a broad overview of the different sections of the probes, but will be broken down further, in the sections of this chapter.

Table 4.1: The author’s intended purpose for the probe sets: original study

PROBE (CRM factor)	**1. Knowledge?**	**2. Purpose?**	**3. Kind of student?**
Pre-test MCQ 1 – 10 (See details appendix b) (C)	Baseline prior knowledge of 5 key concepts in sound embodied in the show (see 3.4.4)	Establish a baseline or starting point for designing show or teaching sound by assessing student prior knowledge. To compare and contrast the three student groups in terms of prior knowledge. For science centre staff or teachers	All three groups (urban, township and rural – U, T and R)
Pre-test MCQ 11 – 13 Attitudinal questions (See details appendix b)	Student attitudes towards school science	Assess initial motivation and attitude of students as a factor affecting learning. For science centre staff or teachers	All three groups

Post-test MCQ 1 – 10 (See details appendix b) (C and R-C)	Assess learning during show in terms of increase in knowledge of 5 key concepts in sound embodied in the show	Measure what has been learnt during the show, and how this differs for the three groups For SC staff to adapt show for different groups and to enhance learning. To motivate funders for funds on account of gains. To encourage visits to the Centre.	All three groups
Post-test MCQ 11 – 13 Attitudinal questions (See details appendix b)	Student attitudes towards the science show	Assess motivation and attitude of students towards the show as a factor affecting learning. For science centre staff.	All three groups
Open-ended Survey Q 1 – 9 Written Explanations (Post-test only) (See details appendix c) (R-C and R-M)	Student conceptions of sound and sound waves. (see 3.4.3)	Probe more deeply into student conceptions to assist in designing show or teaching sound. To compare and contrast the three student groups. For SC staff or teachers	All three groups
Open-ended Survey Q 1 – 7, 9 Drawings (Post-test only) (See details appendix c) (R-C and R-M)	Student visualization of the concepts of sound and waves. (their mental models) (see 3.4.3)	Try to see sound concepts through the eyes of (different) students for SC staff to design show elements to connect with students. Also useful for teachers.	All three groups
Open-ended Survey Q 5 - 9 Select an Answer (Post-test only) (See details appendix c) (R-C and R-M)	Student conceptions of sound and waves. (see 3.4.3)	Probe more deeply into student conceptions by assisting them with suggested answers. To compare and contrast the three student groups For SC staff or teachers	All three groups
Open-ended Survey Q 1 - 9 Source of Knowledge (Post-test only) (See details appendix c) (C-M)	Source of student knowledge in answering questions (see 3.4.3)	Assess student perceptions of whether knowledge was known prior to the show or gained during the show. To assist SC staff in evaluating show.	All three groups
English Language Test MC Q's 1 – 50 (Post-test only) (See details appendix d) (C)	Basic English language ability in terms of grammar, vocabulary and comprehension skills.	Assess student English language ability for SC staff to adapt (or translate) show for different groups. Also to assess validity of other probes in terms of English requirement.	All three groups

This table will again be presented at the end of this section (table 4.4) but indicating how effective each probe was felt to be in terms of achieving its intended purpose in the light of the student data.

4.1 Validation of the pre/post-test multiple choice instrument (Appendix b)

To confirm that these MCQ probes are valid for the intended use and audiences, an item analysis was performed on results for the ten questions, as detailed in section 3.5.2. The item difficulties (p) and discrimination indices (DI) are shown in table 4.2 for both the pre-test and the post-test.

Table 4.2: Item analysis of multiple choice questions (*original* study)

MCQ	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	Ave.
ALL SCHOOLS											
Difficulty p: pre-test	0.25	0.15	0.03	0.29	0.36	0.29	0.26	0.19	0.19	0.23	0.22
Difficulty p: post-test	0.56	0.48	0.06	0.44	0.30	0.54	0.49	0.54	0.45	0.32	0.42
Discr Index PREPOST	0.57	0.77	0.00	0.80	0.57	0.80	0.77	0.83	0.77	0.53	0.64
URBAN SCHOOL											
Difficulty p: pre-test	0.18	0.03	0.00	0.58	0.68	0.35	0.38	0.23	0.23	0.38	0.30
Difficulty p: post-test	0.43	0.73	0.03	0.90	0.53	0.80	0.58	0.83	0.63	0.45	0.59
Discr Index PREPOST	0.7	0.6	0	0.7	0.4	1	1	1	0.9	0.9	0.72
TOWNSHIP SCHOOL											
Difficulty p: pre-test	0.20	0.20	0.03	0.23	0.18	0.13	0.15	0.23	0.23	0.20	0.18
Difficulty p: post-test	0.65	0.43	0.03	0.18	0.18	0.60	0.50	0.53	0.45	0.28	0.38
Discr Index	0.8	0.7	0	0.4	0.3	1	0.9	0.7	0.8	0.3	0.59

PREPOST											
RURAL SCHOOL											
Difficulty p: pre-test	0.38	0.22	0.05	0.05	0.22	0.41	0.27	0.11	0.11	0.11	0.18
Difficulty p: post-test	0.59	0.27	0.14	0.22	0.19	0.19	0.38	0.24	0.27	0.24	0.27
Discr Index PREPOST	0.7	0.3	0.2	0.3	0.3	-0.1	0.7	0.6	0.3	0.4	0.37

When one looks at p-values in table 4.2 for the pre-test for all schools, they were all low with an average of just 0.22. This is very close to the chance value of 0.25 which one would obtain for students just guessing. As the students were naïve and unschooled (in sound and waves), this figure agrees with what one would expect, showing that the results of the questions are a good reflection of the students' *lack of* knowledge in this area. The post-test results are all significantly higher (with the exception of questions 3 and 5) with an average of almost double at 0.42, again as expected, after the students have received instruction through the show. This suggests that the instrument was valid in measuring the 5 basic concepts in sound (see 3.4.4), which were covered during the show, as results were all better post-instruction (except for a slight drop for question 5, and of course question 3, which was not taught on the show and which is discussed further below).

While these p-values are low, it must be kept in mind that the study group deliberately involved naïve, unschooled students, in order to isolate the effect of the show without background interference from their school. If the average p-value across all schools had been > 0.5, I would have been concerned that the questions were too easy. The average post-test value (p = 0.42) is within the acceptable range (0.4 < p < 0.6) for a norm referenced test, where I am seeking to compare performances relative to other students and other groups. The questions yielded a good spread of results both across all three groups (standard deviation on post-test = 21.7 %) and within each group (see section 5.2), also suggesting good validity. Finally, the decrease in average p-value moving from urban to township to rural, is consistent with the decreasing resources, facilities and teachers available to each group, (See student context section 3.1.2) indicating that the questions are a good measure of general knowledge of the phenomenon of sound.

The Discrimination Index (DI) was calculated from the scores for the top 25% of all students in the post-test, minus the scores of the bottom 25% in the pre-test, divided by the number of students in that 25 % group, in this case 30 students. As I was trying to compare and contrast different school groups, the same test had to be used for all three groups, and I needed to look at the DI for the combined group, not individual schools. The data is presented in table 4.2 above and visualized graphically in Fig. 4.1 below.

As shown in Fig. 4.1, with the exception of the comparison question, 3 (see section 3.4.4), these values were all above 0.51, showing good discriminating potential. Six questions showed excellent discriminating potentials with DI's above 0.71. There were no flawed items with negative DI's and the lowest DI (apart from q 3) was 0.53. Hence I can confidently claim that these 10 questions show good discriminating potential when applied to this combined group from the DI-values.

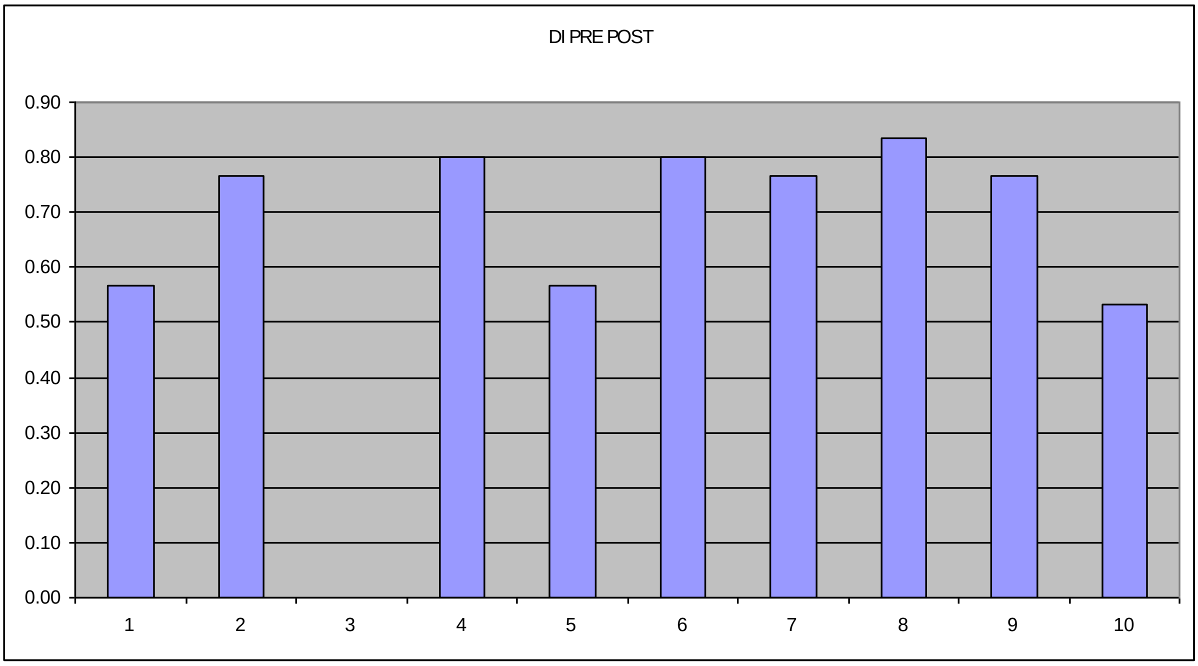


Figure 4.1: Graph of Discrimination Index per question (all schools).

When considering individual school groups (See table 4.2), DI values were not consistent across all three of these, indicating that the questions may work better for some groups than others. The graph below (Figure 4.2) shows that DI values were excellent for the urban group (Average DI = 0.72), good for the township group (Average DI = 0.59) and medium for the rural group (Average DI = 0.37). The overall weakness of the rural group in all areas – especially language (see 5.1.3.1) meant that many students were unable to understand the questions properly and were most likely just guessing, causing results which were randomly distributed and failed to discriminate well between top and bottom students. Thus the DI results suggest that the MCQ instrument is perhaps less suitable for the rural group. Since in the present study, a major goal was to compare and contrast the student groups, I decided to still use the MCQ instrument for the rural group, while bearing in mind the limitations of the data that this instrument would yield for this particular student group. .

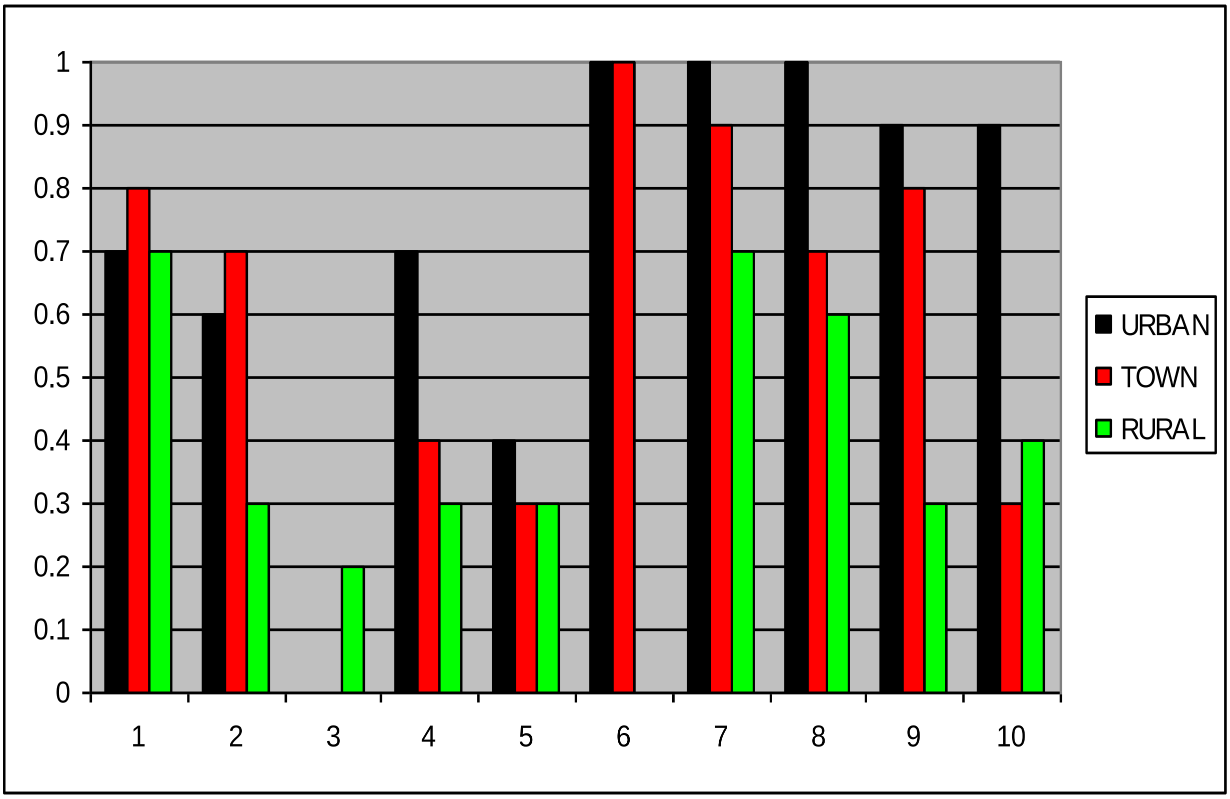


Figure 4.2: Graph of Discrimination Index per question, in school groups

Neither p nor DI considers the performance of the incorrect response options, or distracters. A distracter analysis was performed on questions 1 – 10 for both the pre-test and the post-test, showing the percentage of students choosing each MC option A to D. The analysis was done for all schools together, rather than individually, as the same MC test needed to be applied to all three groups to allow me to compare and contrast their performance. In the graph below (Figure 4.3), the key (or correct option) is coloured red.

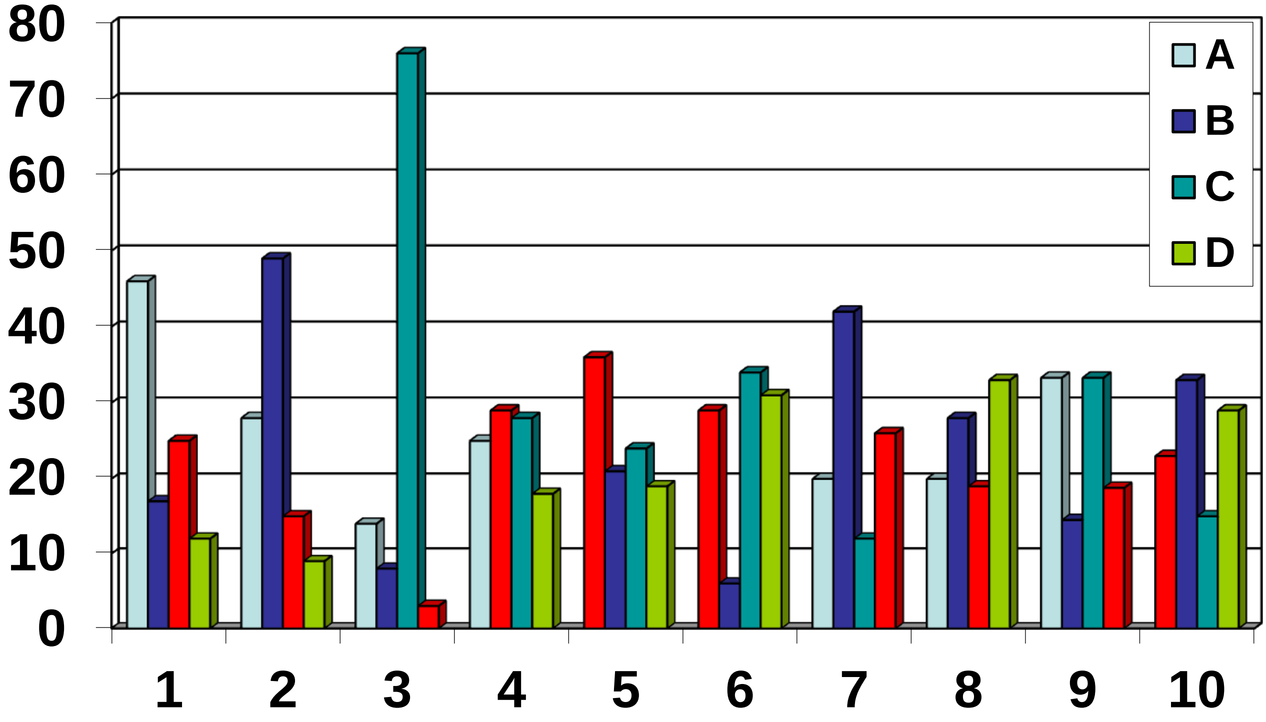


Figure 4.3: Graph of Percentage of students choosing each multiple choice option (A, B, C, D) in the pre-test (all schools – key is coloured red)

In a general sense, the graph above indicates that the distracters worked well as there was a good spread of choices across most of the options with no option receiving no choices, and very few receiving substantially fewer choices than others. This spread may also indicate though, that students are simply unfamiliar with the terms and don't know the answers, so are basically just guessing. Questions 1 – 3 show evidence of a high incidence of prior student conceptions which will be dealt with in the next section (Chapter 5).

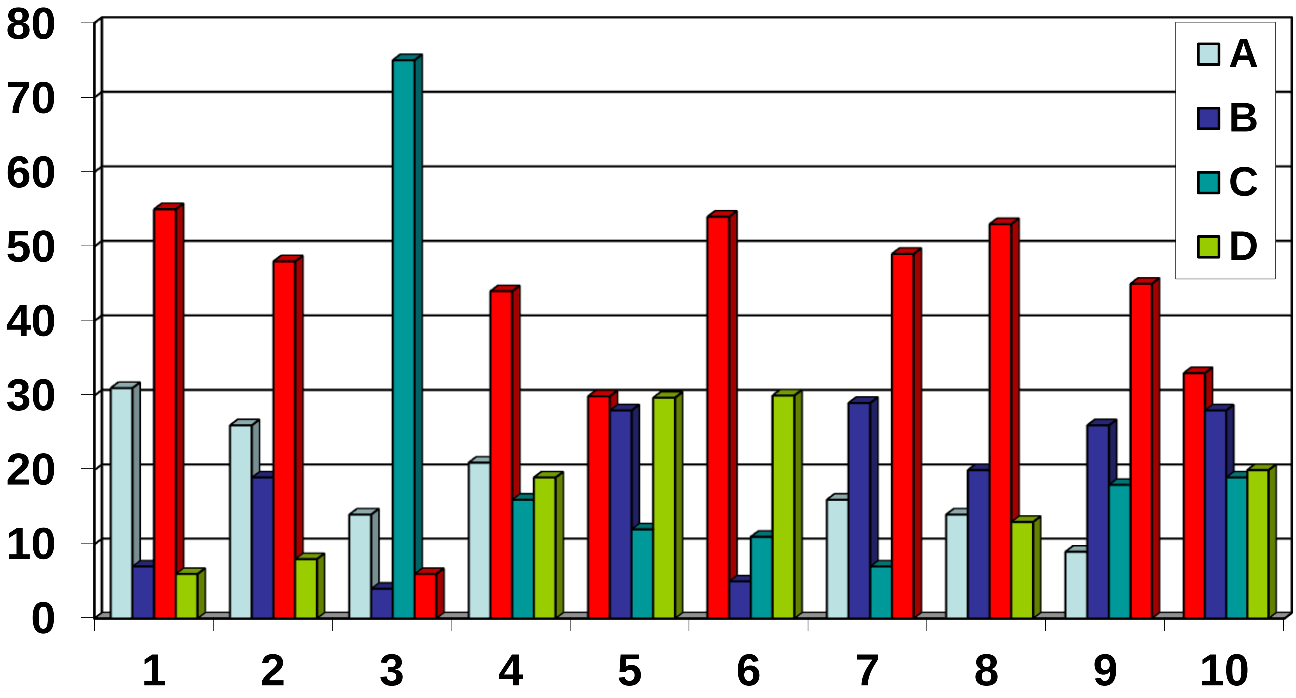


Figure 4.4: **Graph of Percentage of students choosing each multiple choice option in the post-test (all schools - key is coloured red)**

The post-test distracter analysis (Figure 4.4) also showed a good spread over the distracters, but showed moreover that most students had come to the correct selection in every question except the comparison question – 3 (outlined in section 3.4.4). This meant that the key was the most chosen option, and that percentages choosing the distracters were consequently lower. Nevertheless they were still well spread. The options 6(b) and 7(c) were again unpopular choices, and will be discussed below. Questions 9 and 10 yielded interesting answers which contrasted for the different groups (individual group data not shown here), and are dealt with specifically in section 5.2.

Question 6(b) and 7(c) showed evidence of being bad distracters as they were not commonly chosen in either the pre-test or the post-test, although neither was completely ignored. In both cases the options used words (“period” and “interference”) which were probably not familiar to the students, and were hence avoided. The other three options in the two questions were either the same (amplitude and frequency) in both questions, or similar (wavelength and wave-speed) and students probably avoided the options which seemed like “odd words out” or words with which they were unfamiliar (especially given their poor English abilities). It is difficult choosing distracters for naïve, unschooled students as there are a limited number of terms they know. In future these two questions could probably benefit from having a better fourth option. Nevertheless despite these bad distracters, the questions still produced extremely high DI values (0.80 and 0.77) so are still valid for use in the results section.

In addition, if one breaks down the distracter analysis further: although option (b) was unpopular for question 6 in the post-test, it was nevertheless chosen only by students in the bottom third (5 students) and never by students in the top third of the combined group, so still performed satisfactorily. Question 7 (c) was chosen twice by students in the top third, and four times by students in the bottom third, so was probably not a particularly good distracter, and should be replaced in future studies in the light of what this data has revealed.

In conclusion then, the statistical measures indicate that my MC questions are indeed valid for comparing and contrasting the performances of the urban and township student groups but that this instrument should be used with caution when applying it to the rural group. The multiple choice format is known to not be the best way to bring out student conceptions as it is fairly rigid and not very flexible in terms of allowing students to reveal all their scientific conceptions (Schönborn and Anderson, 2008), and this was demonstrated in the present study. However it did provide an efficient means of rapidly screening student knowledge of the key concepts of sound among visitors to the Centre. This single-tier test really evaluates only the “descriptive knowledge” of the participants (Caleon 2010), and is limited in that way. Nevertheless comparison of the pre- and post-test scores showed very encouraging evidence of students’ learning gains across all questions – except question 3. The fact that no learning gains were detected with this question – the content of which was *deliberately* not covered in the show – suggests strongly that this probe is valid in investigating whether I can tie the learning achieved to the show. Question 3’s results, along with the improvements on the other questions where content *was* covered during the show, convince me that the 10 MC questions do actually assess the content in the show, and especially the 5 stated intentions for the show (3.2) and not something else, hence showing construct validity.

In addition to this evidence from the data, multiple choice questions were thoroughly discussed with a panel of 4 experts who confirmed that the questions and answers (underlined in appendix b) are correct and that they do cover the 5 key concepts presented in the show. This panel comprised two experts in science education and assessment, and two experts in Physics. In addition, presentations and informal papers covering this work were presented at 3 national Physics conferences and reviewed by colleagues in Physics education who gave input and made suggestions.

4.2 Development and Validation of the open-ended survey (Appendix c)

As discussed earlier, the open-ended survey was administered *only as a post-test*. The 9 questions on the survey (appendix c), requested students to answer basic conceptual questions that probed both their knowledge and mental models of sound and waves, by:

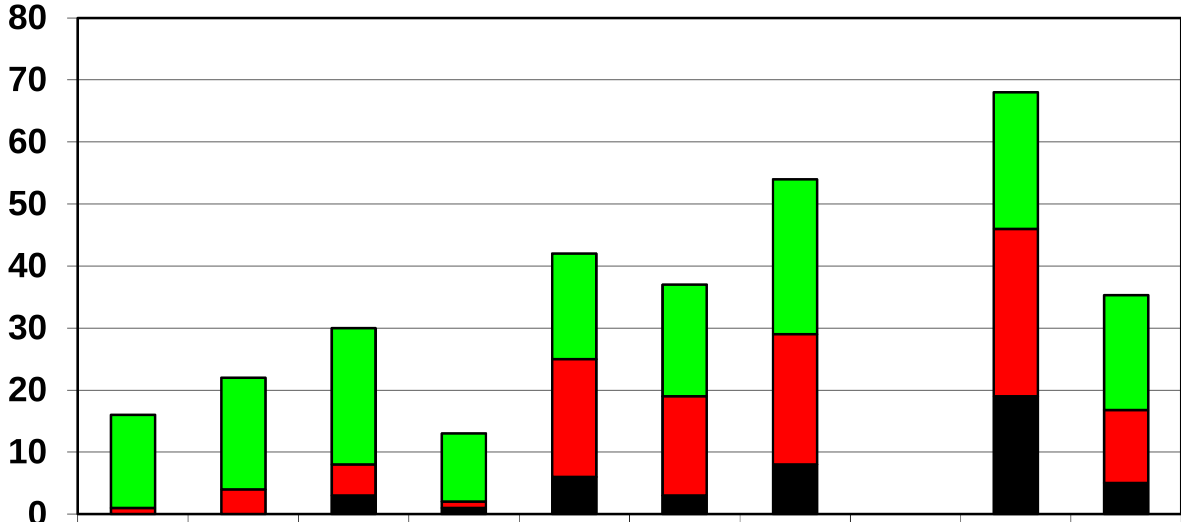
a) Making a drawing or diagram to explain their answer (8 questions)
b) Choosing an opinion from four offered (5 questions)
c) Giving a written explanation (all 9)
d) Indicating the source of their knowledge in answering (i.e. pre-show or after-show) (all 9)

The development of the open-ended survey included extensive testing of the questions in two pilot studies (described in sections 3.4.1 and 3.4.2), in collaboration with a panel of experts, as mentioned in section 4.1 above. Data gathered from these pilot studies led to appropriate modifications to the questions in terms of their content, language and sequence in the test instrument.. As discussed below, I made several important improvements, but as will become apparent in the final results, the final survey was not perfect and could still be modified and improved for future use. Details and comments are given below.

4.2.1 Evaluation and validation of the drawing tasks:

While the drawings provided very interesting and rich data (See results chapter 5) students clearly struggled greatly to produce them. The problem emerged, especially with questions 6 and 9, that students tended to produce drawings which were more *illustrations* than *explanations.* Rural students especially appeared very hesitant to provide drawings and on average exactly half of the students (18.5 out of 37) failed to provide a drawing for *each* question. Township students were similarly loath to submit drawings, and on average about a third of them (12 out of 40) failed to submit drawings for each question. Urban students seemed quite comfortable with drawings and only an eighth of them (5 out of 40) failed to submit drawings for each question.

This was unexpected by the author, as one might expect students who struggle with language (like the rural students) to be *more* eager to draw than to write? Perhaps the difficulty was in understanding the request to draw, and maybe an example should have been given for students in the first question? The number of drawings left blank (out of a total of 117 students) is reflected in the graph below (Figure 4.5) (note that question 8 did not ask for a drawing).



	1	2	3	4	5	6	7	8	9	AVE
RURAL	15	18	22	11	17	18	25	0	22	18.5
TOWN	1	4	5	1	19	16	21	0	27	11.8
URBAN	0	0	3	1	6	3	8	0	19	5

Figure 4.5: Graph of Number of Blanks vs question number (all schools divided into groups)

One also notices an interesting trend which I shall call “drawing-fatigue” here. All students finished all questions and all chose multiple choice options in *every* question – including the last one, question 9. Yet as time went on one sees a definite reluctance to produce a drawing, as evidenced by the increasing number of blank drawings. This suggests that making a drawing requires an effort and as students get more and more tired they are less and less inclined to do so (remembering that at the time of writing question 9 they were at the end of a fairly intensive 2 hour programme). Graphing the total number of blank submissions against question number (which is effectively time or tiredness) one can obtain a fairly good linear fit as below (Figure 4.6):

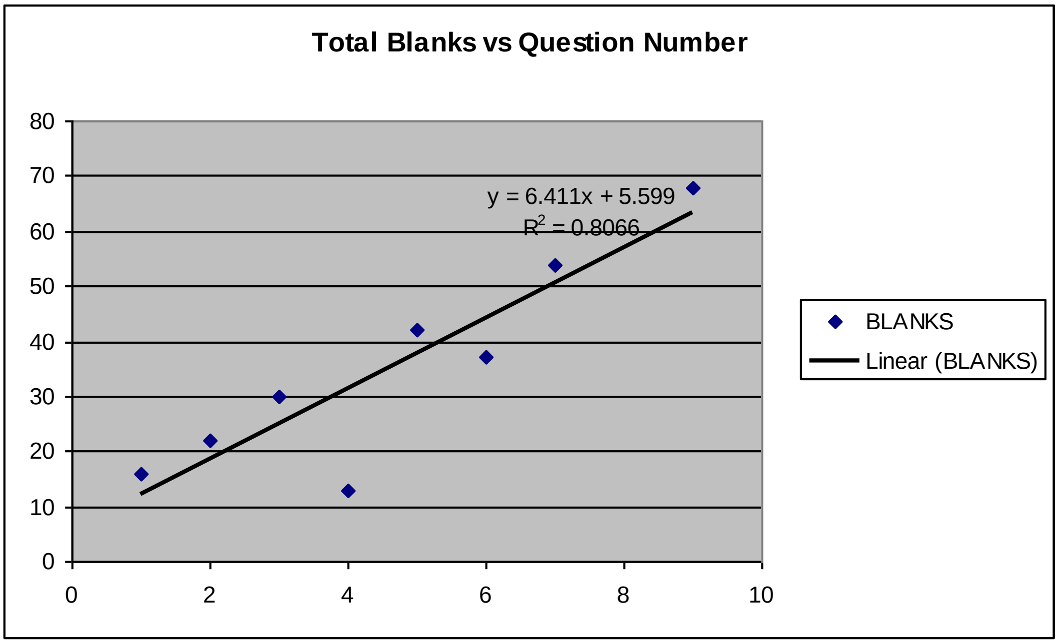


Figure 4.6: Graph of Total Number of blanks vs. question number (all schools)

A study of the number of blank submissions (Fig. 4.6) per question is interesting:

- Questions 1, 2 and 3 formed a conceptual progression requiring students to illustrate how sound waves travel
    1. Through space
    2. Through a wall
    3. Around a corner

  It appears that as the complexity increased, so fewer students were able (or willing) to attempt a drawing.
- <u>Question 4</u> required students to draw a picture almost identical to what they had seen on screen from the data-studio programme. As they had seen this first-hand, they appeared more able to visualize and consequently were more prepared to attempt the drawing, hence fewer blanks.
- Question 5 required students to draw sound waves travelling through different phases of matter, which students clearly struggled to do.
- <u>Question 6</u> asked students to explain why a guitar has a sound box. Most were happy just to draw a guitar (an example of *illustration* rather than the required *explanation*) so felt fairly confident to attempt this – so the number of blanks dropped.
- Question 7 required students to show sound waves interacting with the ear, again a more complex extension of question 1's drawing, and that difficulty and fatigue showed in the increasing number of blanks

- (Question 8 did not ask for a drawing)
- Question 9's drawing was difficult to visualize but also at the end of a long survey, and students were clearly anxious to get finished, so the number of blanks increased accordingly, to a rather alarming 68, or more than half the surveys .

Even with the anomalies of questions 4 and 6, the fit is still good with a regression coefficient of 0.81, which seems to indicate that fatigue as the survey goes on results in an ever increasing number of blanks. This "fatigue" was also observed by Wittman et al. (2003):

> "… the high number of blank responses on the candle flame response may be due to the candle flame question coming in the later half of the pre-test."

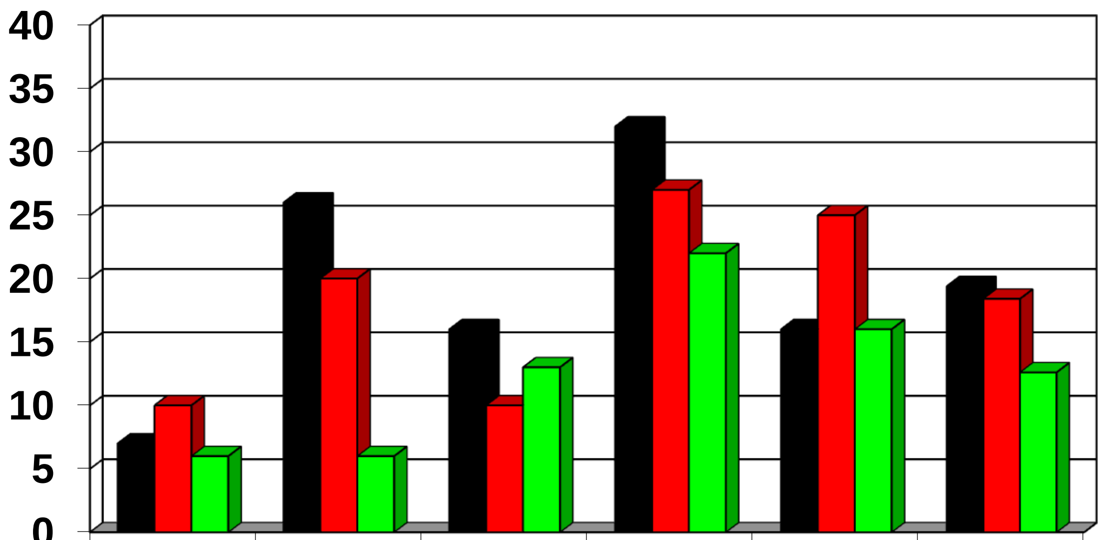
This high correlation shows that this factor is clearly evident and one needs to find a way to mitigate against it. Perhaps the questions requiring drawings should all be at the start of the survey, or perhaps one should not accept any submissions where students have not at least *attempted* the drawing. As I was taking in over 1000 sheets of paper, this would have been very difficult to do in the time available. It would appear that students are not used to producing drawings or diagrams, so it is questionable how effective this probe is if the activity is so new to them. In addition, it was difficult to use this probe to compare the groups, as the number attempted from each group was so different.

It was hoped to be able to compare the information received in the drawings with written answers and multiple choices. Unfortunately the large number of questions left blank, and the extremely poor quality of the written answers made it very difficult to do so. Most drawings did agree with written answers, but more in the sense of their being *illustrations* than further *explanations*. (Written answers also were often more *descriptions* than *explanations,* echoing Cooke and Breedin's (1994) finding).

The drawings were validated by the same panel of 4 experts mentioned previously to ensure that each probe does indeed test what I expect it to probe. In addition the "model answer" drawings against which student answers were measured (shown in appendix j) were validated by these four, and reviewed at Physics conferences. The detailed results of the pilot study (appendix m) and the open-ended survey pilot (appendix n) contain many examples of both sound and unsound drawings provided by students, validating these questions in terms of their ability to elicit a range of responses from the participants. Further examples of both sound and unsound drawings are shown in section 5.2.

4.2.2 Evaluation and validation of the “Choose an Opinion” tasks

Questions 5 – 9 of the open-ended survey asked students to identify with one of four opinions offered by cartoon characters: - essentially a five-question, four-option, multiple choice post-test. Whereas the drawings required students to reveal their mental models of sound, here I gave them models and asked the students to interpret them. Results for the different school groups per question are shown below in Figure 4.7. These results mirror those in the multiple choice post-test, suggesting good agreement across probes.



	5	6	7	8	9	AVE
URB	7	26	16	32	16	19.4
TOWN	10	20	10	27	25	18.4
RUR	6	6	13	22	16	12.6

Figure 4.7: Graph of Number of correct answers per question (OES 5 – 9), in school groups:

Besides validation by the panel of experts mentioned before as to what the author predicted the purpose of each probe was, probe validity was also determined by calculating the:

- item difficulty index (p) of each item
- item discrimination index (DI) of each item

and performing a distracter analysis of each item. The results are shown in the table below (table 4.3).

Table 4.3: Item analysis per open-ended survey question 5 – 9 (all schools)

QUESTION	5	6	7	8	9	AVERAGE
p-VALUE	0.20	0.44	0.33	0.69	0.49	0.43
DISCRIM INDEX (DI)	0.37	0.60	0.53	0.77	0.70	0.59

The item-difficulty indices (p-values) are reasonable for a post-test for these groups of students and the average difficulty index (0.43) for these 5 questions was almost identical to that for the MC post-test (0.42), showing good agreement across different probes, despite having quite different questions.

With the exception of question 5 (comparison question) the DI values are all either good or excellent, showing that these questions are valid in terms of their ability to discriminate between the top and the bottom students. Even including q 5, the average DI of 0.59 is good.

A graph of student choices across all three groups for the distracter analysis is shown below (Fig 4.8):

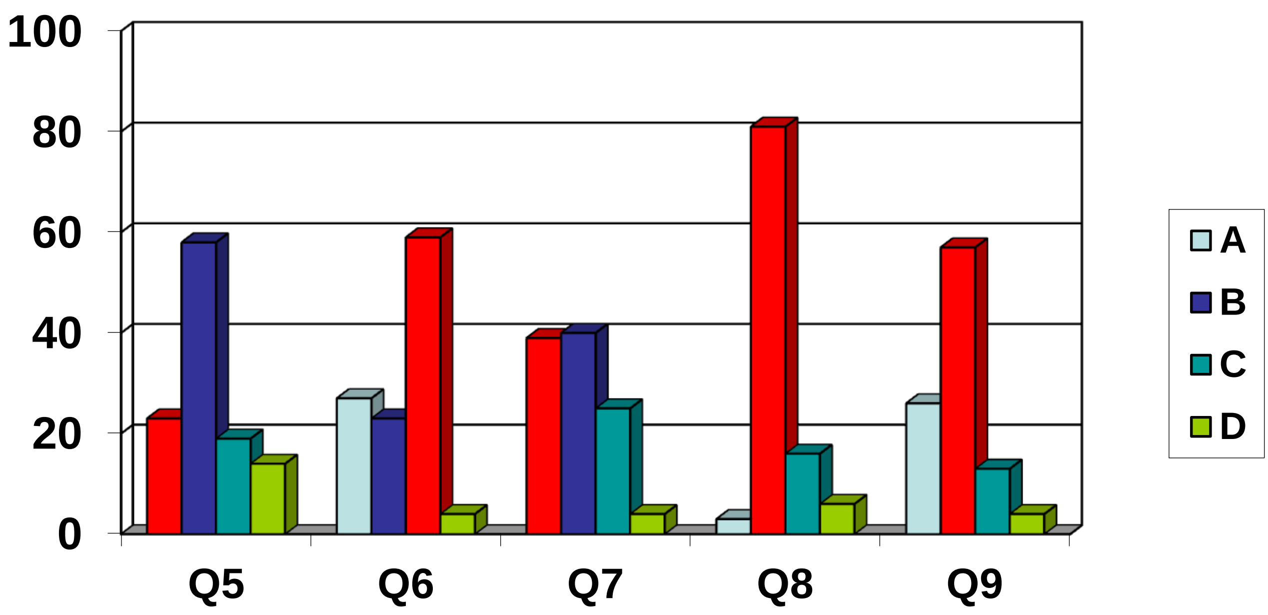


Figure 4.8: Graph of number of students choosing each option OES questions 5 – 9 (all schools- key is coloured red)

One disappointment was that very few students chose option (d) ("I don't agree with any of you"), on average only 6 students per question, or just 5 % of the 117 students. This may represent lack of confidence to venture into the unknown, but in fairness option (d) was never the correct one, and having even one question where (d) was correct might have changed this trend. Students perhaps also felt that they may have to explain more in the written section if they answered (d). As students didn't really consider this option, the choices were effectively limited to three which was not ideal. When looking at student choices in the graph above, and considering that these questions were administered as a post-test only, the results are very similar to the multiple choice post-test. Question 5 is effectively a repeat of MCQ 3 ("Where does sound travel the fastest?") and the same student conception ("B: In gases") is evident here, as this concept was deliberately not covered in the show. Again the correct answer ("A: In solids") was chosen by few students (23 out of 117). This shows good triangulation with the multiple choice probe.

In questions 6, 8 and 9 the correct answer emerges quite clearly as the one most chosen. Leaving out option (d) ("I don't agree with any of you": see discussion above) the other two distracters were rarely chosen, but were never completely ignored. Questions 6 and 8 were directly dealt with in the show with demonstrations and extensive discussion, and the clear choice of the correct option for these questions bears that out. The concept covered by Question 7 ("Why does a loud sound hurt you ears?") was mentioned towards the end of the show, but not at any great length, and the students were consequently less clear about the correct answer. In retrospect the correct answers for questions 8 and 9 could have been better stated (as discussed more fully in appendix j), and should be refined before the questions are used in future studies. The fact that many students nevertheless chose the correct answer points to these questions (especially q 8) leading directly on from information delivered in the show.

Apart from option (d) discussed above, the distracters seemed to work well and there were none which were never chosen. Lessons learnt from the Open-ended Survey Pilot (see 3.4.2) helped to exclude some weak distracters before this final survey. Question 8 (a) relates to a comment ("men have deeper voices than ladies because they are taller") discussed at length during the show, so the fact that very few chose it probably indicates that it was identified by most as a spurious explanation.

This section seemed to be an effective probe in testing student conceptions after the show, and students very rarely left it blank and almost always were willing at least to attempt this section. The correct answer mostly emerged as the most popular one, and prior conceptions appeared to have been dealt with during the show. The fact that q 5 showed a low p-value (and less than chance) agrees with MCQ 3 results and again suggests that these questions are valid in indicating that learning can be tied to the show.

The problems with this probe were more in terms of how it related to the next section (Written Answers). This method of questioning was used to try to assist students by asking them to identify with a friend's explanation, rather than come up with it from scratch, as discussed in 3.4.2. While this probably did assist students, many of them simply rewrote what was in the voice bubbles in the next section, rather than coming up with an elaboration or further explanation of their own. This was especially predominant in the rural students, and to some extent in the township students, where language skills were weak. In retrospect I do not think this was a good method to use (together with asking for written answers) with groups who have poor language skills, as they mostly just repeat what is given, and one does not get the chance to explore their own ideas. In future studies I would recommend keeping multiple choice and written answers separate. Hence the intention of the questions was to give the students models and ask them to interpret them. Instead the students merely reproduced my models, or rewrote my words, which didn't help to reveal their mental models.

Again these questions were the subject of intense discussion and changes detailed in section 3.4.2. The open-ended survey pilot was limited in that it was applied only to an urban group and some of the language issues mentioned above did not emerge. Nevertheless expert validation was done on these 5 questions and their answers before they were used.

4.2.3 Evaluation and validation of written answers:

Written answers were very disappointing generally and language was clearly a major issue.

> "We showed that language may be a significant barrier in effective communication related to sound." (Hrepic 2010)

Students seemed quite unable to express answers clearly, even though it seemed at times that they had an idea of what was going on (or chose the correct answer in the multiple choice). They struggled to form models to explain what was happening by themselves. Hrepic (2010) in a study of *English-speaking university students* even reported:

> "… students rarely displayed coherent reasoning at the level of a structured model without being prompted with additional questions, especially in preinstruction interviews. For example, models were rarely found simply after the first general question about propagation of the speaker's voice. However, they were frequently found when additional model targeting questions in the context 1 and 1a were asked. Thus, some of these models may have been generated on the spot in the student's attempt to provide some rationale for the presented situations. So although sound is one of the most common of daily-life phenomena, students seem unlikely to form a mental model of the phenomenon unless they are requested to provide some explanation for it."

In the light of his findings, and in retrospect, it was perhaps expecting too much to hope that Grade 9 students with weak language ability would have the science, language, concepts and indeed confidence to give full and correct answers to questions. Furthermore, Hrepic's comments above indicate that students may need to be helped to form models by additional questions and probing. As I did not have the opportunity to conduct follow up interviews, there was no chance for students to display this "coherent reasoning" or to present more carefully structured models. Clearly this is too much to hope for in a short science show. Nevertheless the answers given were illuminating:

Because of this concern, a simple English language proficiency test was administered to about half of the students in each group 6 months after the main research was done. The results of this test are in Chapter 5 and bore out my concerns about language – especially with the rural group.

As an example of weak English skills and subsequent misunderstanding of a question, consider Open-ended Survey (OES) question 3: A small change in layout compared with OES 1 and 2 led to the majority of students misunderstanding this question. Whereas questions 1 and 2 asked: "explain how sound travels from A to B", this question said: "Explain to me how it is that B can hear A around a corner?" and then asked: "Explain to your friend how this is possible using a drawing, and in words."

Students focused on how it is *possible* for B to hear A, and (understandably) answered – "…because the sound is very loud." They were reflecting on their own life experience (shout loud if you are far away or round a corner) rather than performing the unusual abstraction of drawing and explaining sound waves. In total, 36 of the 111 students who answered mentioned that the sound must be loud, with 24 of these giving this as their only answer. This question needed to repeat the wording from 1 and 2 and ask: "Explain how sound travels from A to B when they are around the corner from each other."

Another language problem was evident in OES 6, this time in student answers rather than in the question itself, but the question could have been more carefully stated. It seems that most students understood that the box amplifies the sound, but (judging from the full written answers) many understood "higher"" to mean "higher volume" or louder, and "clearer" to mean "more clearly heard" or louder. Thus for many students there were no real options as A, B and C all said "louder" to them in different words. This is one of the problems of a written survey over an interview, where one cannot ensure that the question is understood. If this question were reused in future, I would write:

- A) The box is to make the sounds more high-pitched (or higher in pitch)
- B) The box is to make the sounds more distinct, or clear.

The confusion between pitch (high and low) and volume (loud and soft) has been a feature of most questions in this survey and in the literature generally (item 8 in my difficulties table) and is something worth taking note of in order for this probe to be more effective.

Sadly this section was not an effective probe for the rural students. Their language skills were so weak that their answers were on the whole almost unintelligible. Once one had excluded those who didn't attempt a written answer and those who simply rewrote the question (or what appeared in the voice bubbles) there was only a handful (generally less than 5) of useful answers. In future studies it would be wise to do a basic English language test first, and then to use those results to assess the suitability of further probes for each group. Thus clearly the normal process of probe validation was difficult for these types of probes as the probes were clearly not valid for most students in my target sample and would need to be extensively modified, perhaps even translated into isiZulu before they could be useful. This doesn't, of course mean the probes won't be useful and valid for different sample of students perhaps with English as first language.

The questions requiring written answers were validated by the same panel of 4 experts mentioned previously to ensure that each probe does indeed test what I expect it to probe. In addition the "model answers" against which student answers were measured (shown in appendix j) were validated by these four and discussed with colleagues at Physics conferences. Samples of sound and unsound student answers are shown below to confirm that the questions are valid in terms of being able to elicit a range of answers from sound to unsound. Urban written answers were mostly much better than those of the other groups (especially rural) but these questions showed validity in that they were able to elicit a range of answers from all the different school groups. Answers were of course related to their choice of suggested answer ("choose an opinion") which they were asked to further explain in detail for the written answer.

Question 1: Explain how sound travels from A's mouth to B's ear in the diagram.	
Model answer	"When A speaks, he produces vibrations in the air, which travel away from him as a longitudinal (sound) wave. When this wave enters B's ear it causes his ear-drum to vibrate, which allows him to hear the sound."
Sound answer	(U) "because of sound waves that are produced and therefore travel towards the ear and are received by the ear drum and the sound is registered to the brain for understanding"
Unsound answer	(T) "A have to use frequency (Pitch) so that B can hear what he is saying"

The sound answer is imperfect but pleasing and on the right track. The unsound answer shows confusion between waves, frequency, pitch and volume.

Question 2: Explain how sound travels from A to B through a wall.	
Model	"When A speaks, he produces vibrations in the air, which travel away from him as a longitudinal (sound) wave and spread out in all directions. When these waves strike the wall in front of him, they cause it to vibrate very slightly and these vibrations pass through the wall. Inside the room these small vibrations in the wall cause the air to vibrate and a wave spreads out through the air traveling away from the wall. When this wave enters B's ear it causes his ear-drum to vibrate, which allows him to hear the sound."
Sound	(U) "The Sound travels from person A's mouth and it vibrates through the wall, transmitting sound to person B."
Unsound	(R) "because the sound will come over in a brick wall and a person outside will speak aloud so the sound will pass through"

The sound answer is perhaps a little incomplete, but is on the right track. The unsound answer shows confusion between whether the sound goes over the wall or through the wall.

Question 3: Explain how sound travels from A to B around a corner	
Model	"When A speaks, he produces vibrations in the air, which travel away from him as a longitudinal (sound) wave and spread out in all directions, but are blocked (or at least severely damped) by the wall at his side. At the corner of the wall the waves are able to spread out around the wall and travel to B. When this wave enters B's ear it causes his ear-drum to vibrate, which allows him to hear the sound."
Sound	(T) "carries on with a widen up sound to an open space."

Unsound	(R) “It can travel in air the position doesn’t change anything the wall is just a shape.”

The sound answer is far from perfect but at least shows some thinking about the sound *spreading out* into the space around the corner. The unsound response shows the student disregarding the effect of the wall on the sound and focussing on sound needing air to travel.

Question 4: Explain the difference between music and noise	
Model	“Music can be represented by a wave which is regular, repetitive and which has a recognisable (smooth) shape. A representation of a sound wave from noise is irregular, non-repetitive and has no recognisable shape: - a jagged shape. “
Sound	(R) “The noise is a mess with no order but music is nice smooth with clear tone.”
Unsound	(T) “The difference between noise and music is that noise has high frequency than music. Music has soft and long frequency wave.”

The sound answer correctly identifies the patterns seen in the two wave-forms, albeit not using the correct terms. The unsound answer wrongly points to frequency as the difference between music and noise.

Question 5: Where does sound travel the fastest?	
Model	“ Sound travels fastest through solids because of the elastic properties of that phase, which allow the vibrations to be efficiently and swiftly transferred from one molecule/ particle to another.”
“Sound”	(R) “In solid because sound need wood to play properly like a guitar.”
Unsound	(U) “B in a gas – the particles in a gas are less dense and they allow sound to penetrate through that particular medium.”

Almost no students answered this comparison question correctly as nothing in the show addressed it. Some linked the question to what they had seen in the show (like the sound box on a guitar) but still did not really answer correctly. Most students answered gases and gave explanations related to “spaces between the particles.”

Question 6: Why does a guitar have a wooden box behind the strings?	
Model	“When the guitar-strings vibrate on their own, they are very thin and don’t move much air. When they vibrate on the box, the whole sound box vibrates, creating a much larger

	vibrating surface and consequently larger amplitude vibrations, giving louder volume."
Sound	(U) "The box is hollow so in that way the sound is louder because the box helps the strings make a loud noise and also the box makes noise because of the hollowness."
Unsound	(R) "the box is to make sound clearer because we cannot hear the sound of the guitar without a solid box like wood."

The sound answer is far from perfect but does bring out the idea of "hollowness" to add to what I saw of the students' mental model. The unsound answer mixes up loudness with clarity ("clearness").

Question 7: Why does a loud sound hurt your ears?	
Model	"A loud sound has very large amplitude vibrations. When it enters the ear, it causes the eardrum to vibrate with large amplitude vibrations – more than it is used to. This causes pain."
Sound	(U) "A loud sound will have large vibrations which may be too large for our ears and therefore it will hurt the sound is too large and also very rough. It is too much for our ear to handle."
Unsound	(T) "It's C because the loud sound is very nice to listen to it but the thing is that it's annoying if it's too loud because it is rough."

The sound answer is on the right track although it has not tied the "large" vibrations to an increase in amplitude. The unsound answer attributes the pain to the "roughness" of the sound.

Question 8: Why do men have deeper voices than ladies?	
Model	"When boys become men, their vocal cords change with puberty (adolescence). They become longer (or) thicker (or) looser. This makes them vibrate more slowly, making their voices lower. This does not happen to women."
Sound	(U) "Men's vocal cords change as they go from boys to men. They cords get thicker and longer, when they turn into men."
Unsound	(R) "Because they are growing up."

The sound answer is pleasing and shows understanding of the root cause of the phenomenon (vocal cords changing). The unsound answer shows a "surface" explanation which does not address the root cause.

Question 9: Can we hear sounds in space?	
Model	"Sound travels in longitudinal waves which require a medium (like air). In space there is no air, only a vacuum, so sound waves can't travel."
Sound	(T) "the sound waves need air to travel and without air they cannot travel so therefore you cannot hear sound from outer-space …"
Unsound	(U) "You can hear sound in space. How loud it is depends on the atmosphere."

The sound answer is pleasing in that the student has picked up on the medium (air) required for sound to travel. The unsound response indicates that the student either misunderstands sound or atmospheres but in either case has not answered correctly.

4.2.4 Open-ended survey: Conclusions:

While the open-ended survey yielded interesting results, it is perhaps ultimately better suited to a classroom situation where instruction is more systematic and students have time to develop their own concepts and visualizations, than to one forty minute show. The test was too long and detailed to be performed as a pre- and post-test, so it was difficult to use it as a tool to measure learning in the show. It would appear to be best suited to looking at student preconceptions and should perhaps in future be administered before the show on a naïve, unschooled group, to provide an understanding of student's base knowledge on which to build the show. Nevertheless it yielded very useful data for answering RQs 1 and 2 b, and in confirming some of RQ 2 c in terms of learning during the show.

English language issues were a major issue, especially for the rural school, and these were not fully appreciated during the Open-ended Survey Pilot which was conducted only with an urban group. These made the written answers almost completely invalid as a tool for evaluating the rural school. The "select an answer" questions (5 – 9) worked well in themselves, but adversely affected the validity of the written answers by allowing weak students merely to rewrite their options.

4.3 Summary of validity:

The table from the start of this chapter is reproduced here. In the light of the discussions above a summary is given as to the perceived effectiveness of each probe in terms of the objectives below, and for each of the school groups. Effectiveness will be indicated using a five point scale:

- very limited
- limited
- average
- good
- excellent

Multiple choice questions worked well generally, although statistical indicators (especially DI) showed that they performed excellently for urban students and not very well for rural students. Open-ended survey written explanations did not perform well at all on the whole and were a reflection of students' extremely weak English language abilities. For rural students answers were so weak as to be almost unusable. Drawings were poorly done by all students and this tool suffered from the fact that students were extremely reluctant even to attempt drawings. The "choose-an-answer" questions which were post-test only performed well on the whole across all groups, although less well for rural students. While performing well they nevertheless caused a problem with the written answers which is discussed above. Source of knowledge questions provided useful information but worked less well for township and rural students who often left them blank or "unsure". Finally the English language test appeared to work well across all groups as the results correlated well with general ability in the other sections.

Table 4.4: Validity of probes for different factors based on analysis of student results: original study

PROBE (CRM factor)	**1. Knowledge?**	**2. Purpose?**	**3. Valid for what kind of student?**
Pre-test MCQ 1 – 10 (See details appendix b) (C)	Baseline prior knowledge of 5 key concepts in sound (see 3.4.4) GOOD	Establish a baseline or starting point for designing show or teaching sound by assessing student prior knowledge. For SC staff or teachers GOOD	U – EXCELLENT T – GOOD R – AVERAGE
Pre-test MCQ 11 – 13 Attitudinal questions (See details appendix b)	Student attitudes towards school science GOOD	Assess initial motivation and attitude of students as a factor affecting learning. For SC staff or teachers GOOD	U – EXCELLENT T – GOOD R – GOOD
Post-test MCQ 1 – 10 (See details appendix b) (C and R-C)	Assess learning during show in terms of increase in knowledge of 5 key concepts in sound EXCELLENT	Measure what has been learnt during the show. For SC staff to adapt show to enhance learning. To motivate funders for funds on account of gains. To encourage visits to the Centre. EXCELLENT	U – EXCELLENT T – GOOD R – AVERAGE
Post-test MCQ 11 – 13 Attitudinal questions (See details appendix b)	Student attitudes towards the science show. GOOD	Assess motivation and attitude of students towards the show as a factor affecting learning. For SC staff. GOOD	U – EXCELLENT T – GOOD R – GOOD (mostly due to English ability)
Open-ended Survey Q 1 – 9 Written Explanations (Post-test only) (See details appendix c) (R-C and R-M)	Student conceptions in sound and waves. AVERAGE	Probe more deeply into student conceptions to assist in designing show or teaching sound. For SC staff or teachers. AVERAGE	U – GOOD T – LIMITED R – VERY LIMITED (mostly due to English ability)

Open-ended Survey Q 1 – 7, 9 Drawings (Post-test only) (See details appendix c) (R-C and R-M)	Student visualization of concepts in sound and waves. GOOD	Try to see sound concepts through the eyes of students for SC staff to design show elements to connect with students. Also useful for teachers. AVERAGE	U – AVERAGE T – LIMITED R – VERY LIMITED (mostly due to unwillingness to attempt drawings)
Open-ended Survey Q 5 - 9 Select an Answer (Post-test only) (See details appendix c) (R-C and R-M)	Student conceptions in sound and waves. GOOD	Probe more deeply into student conceptions by assisting them with suggested answers. For SC staff or teachers GOOD	U – GOOD T – GOOD R – AVERAGE (but affects written answers adversely)
Open-ended Survey Q 1 - 9 Source of Knowledge (Post-test only) (See details appendix c) (C-M)	Source of student knowledge in answering questions GOOD	Assess student perceptions of whether knowledge was known prior to the show or gained during the show. To assist SC staff in evaluating show. GOOD	U – GOOD T – AVERAGE R – AVERAGE (T and R left many blank or unsure)
English Language Test MC Q's 1 – 50 (Post-test only) (See details appendix d) (C)	Basic English language ability in terms of grammar, vocabulary and comprehension skills. GOOD	Assess student English language ability for SC staff to adapt (or translate) show for different groups. Also to assess validity of other probes in terms of English requirement. GOOD	U – GOOD T – GOOD R – GOOD (results correlated well with general ability in other sections when comparing groups)

Finally, as an example of comparison between probes, let me consider 2 of the students' difficulties and the questions which related to them:

Students' difficulty 12 related to the difference between music and noise. This was specifically covered during the show with many demonstrations and computer simulations (see appendix a). It was also probed with MC question 8 and open-ended survey question 4. In the MC pre-test, scores were quite low, with only 19 % of students from all schools getting the correct answer. In the post-test this rose to 53%. In the open-ended survey question 4 (which was only a post-test) a very similar percentage of students (56 %) drew 2 different waves representing noise and sound in the drawing section. While their drawings were not perfect, they reflected the general differences asked for.

Similarly written answers were readily attempted and gave reasonable explanations to explain the question. What was pleasing was that written answer brought out a wide variety of terms for both music and noise, which differed markedly between the three student groups, adding interesting extra information to the MC results. In addition the post-test MC results showed a similar trend to the open-ended survey in that urban results were better than township, which were better than rural in both cases. Finally, when students were asked whether they acquired this knowledge before or during the show, the ratio of during: before was high at 2.26, linking well to the reality of their experience.

In contrast student difficulty 4 relating to the speed of sound was deliberately not covered during the show as a comparison question. It was probed with MCQ 3 and OES Q 5. With the multiple choice it was the only question where there was no improvement pre- to post-test, and the same result for all three student groups. In the open-ended survey question, all groups again answered very poorly as the question had not been dealt with during the show. The "choose an answer" section had the lowest result of all of these questions at less than 20 % answering correctly. Written answers were consequently also weak and incorrect on the whole. The drawings for question 5 were not even attempted by half of the rural and township students and even those who did answer failed to provide good explanations of the phenomenon. Finally the reported ratio of knowledge acquired during the show to before the show was the lowest for any question at 1.14, agreeing with the lack of instruction on this question in the show.

In conclusion then, in both these two test cases where the phenomenon was well covered in the show or not covered at all, I saw triangulation across all of my probes with similar results for the different student groups emerging in each case. This suggests then that these probes are valid across all groups (though not necessarily equally valid as discussed above) and that I can proceed to use them to obtain data for my results section.

Chapter 5: Results and Discussion: Original study (Stage 3)

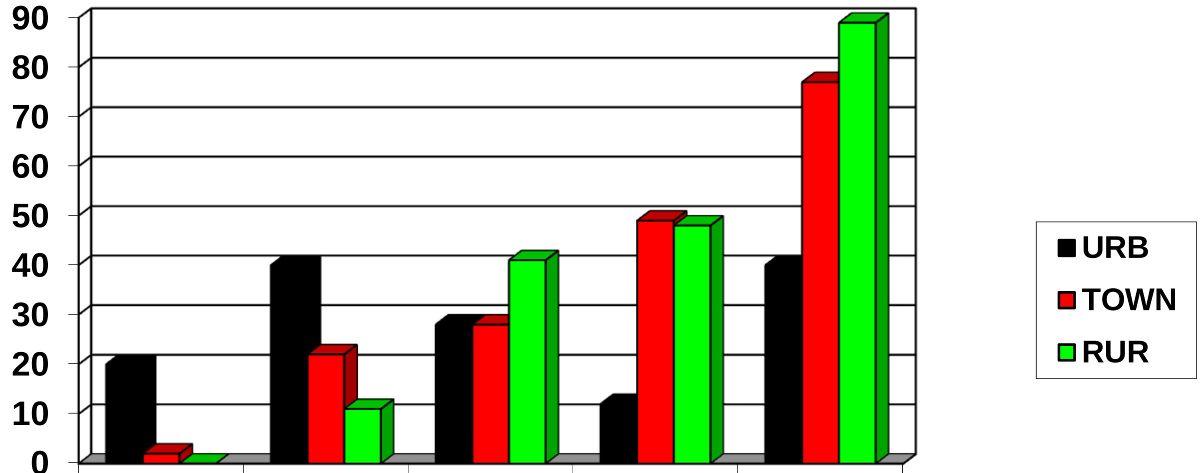
In this chapter I present and discuss the results of the student responses to the instruments developed in chapter 3 and validated in chapter 4, for the original study group. In particular I will address RQ 2, namely: How do the students from the *urban, township and rural* school groups compare, after experiencing *the sound show*, in terms of: a) General attitude (section 5.1); b) Conceptual and visual difficulties with respect to sound (section 5.2) and c) Prior knowledge and learning relating to the show (section 5.3).

This chapter will end with section 5.4 in which I will summarize and compare the general performance of these school groups across the various measures. In writing this chapter, I considered it important to not overemphasize differences between groups in a way which could mask variations within one particular group (Lee 2003). What follows are at best generalizations which point to trends which should always be used as guidelines rather than as prescriptions.

5.1 General attitude of the student groups

Responses to the main attitudinal questions (see details of the instrument in appendix b) are presented in this section in an attempt to understand what motivated the participant groups to visit and participate in the centre activities and how this would relate to their educational and socio-economic backgrounds as discussed in Section 3.1.2. These findings will then be related in the subsequent sections to student performance on the various measures of understanding, difficulties and learning about the phenomenon of sound.

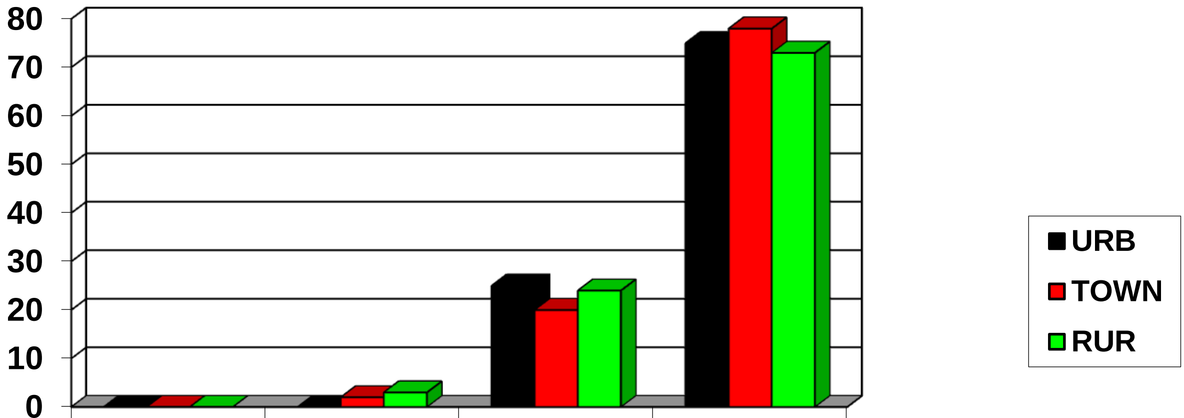
Regarding the question to the students about, “How do students find science classes at school?”: found in MCQ 13e (See appendix b),it was surprising that those students from the more privileged urban school, with better educational backgrounds and opportunities, found science to be boring or just OK (60 % combined for urban students; see Fig. 5.1). In contrast, figure 5.1 shows that more than 75 % of students from the township and rural school groups found science at school to be quite, or very enjoyable. This suggests that students’ positive attitude to the show was not related to their disadvantaged circumstances.



	BORE	OK	Q ENJ	VER ENJ	q or v enj
URB	20	40	28	12	40
TOWN	2	22	28	49	77
RUR	0	11	41	48	89

Figure 5.1: Student enjoyment (%) of science classes at school in school groups.

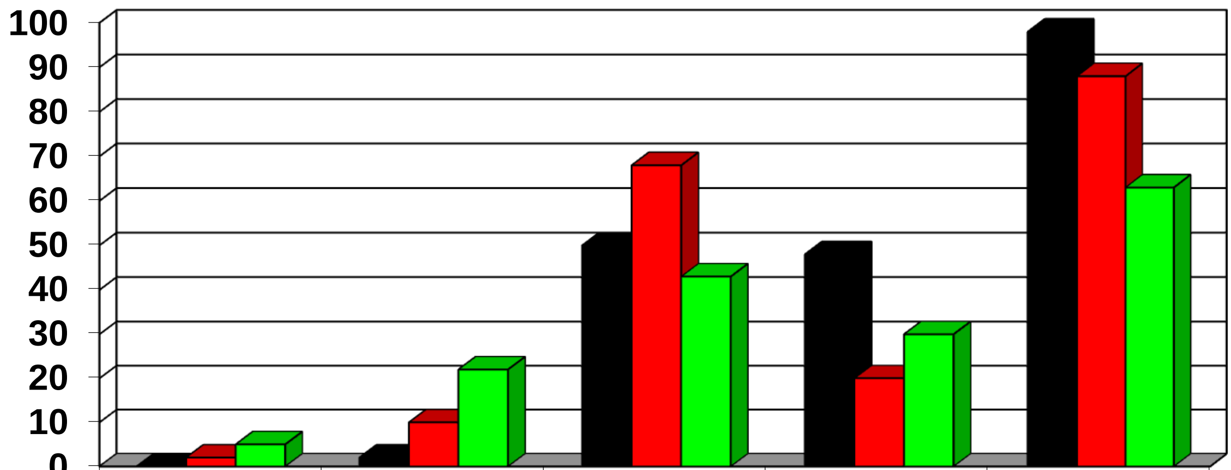
In response to the MCQ about “How enjoyable was today’s science show?” (see MCQ 11p in appendix b and Fig 5.2 below) most students agreed in their responses, that the science show had been quite or very enjoyable especially when compared with MCQ 13e in which the same groups disagreed about how enjoyable school science lessons were. As shown in Fig 5.2 below, I would have expected the more sophisticated and demanding urban students (see section 3.1.2) to rate the show lower (as happened with urban students in Pietermaritzburg in the pilot study – see the detailed summary of these results in appendix m). Perhaps this is due to the fact that most of the urban students were not highly privileged in terms of opportunity to be educated? But it does tell me that the Science Centre can effectively entertain students from all three groups and that enjoyment of the show is not a factor which causes differences between the three groups. This also suggests that the usual post-show evaluation which asks the question: “How much did you enjoy the show?” does not yield very fruitful data. Thus the data showed that most of the visitors I studied *did* enjoy the science-show confirming my past experiences with other groups of students. However, it remained to be seen whether enjoyment by students could be equated to how much they learned during the show. This was subsequently investigated, the results of which are presented in the sections below.



	BORE	OK	Q ENJ	VER ENJ
URB	0	0	25	75
TOWN	0	2	20	78
RUR	0	3	24	73

Fig 5.2: Student enjoyment (%) of science show in school groups

Regarding the question to the students about, “How understandable was today’s show?” (MCQ 12p), although enjoyment was common across all three groups, understanding was not – most probably a function of language. As shown in Fig 5.3 below, the majority of urban (98%) and township (88%) found the show easy or very easy to understand as compared to just 63% for rural (5th column). It shows that I need to focus more clearly on understanding with rural groups, and perhaps worry less about enjoyment as they will probably enjoy the outing no matter what.



	VERY DIFF	DIFFICULT	EASY	VERY EASY	e or v.e.
URB	0	2	50	48	98
TOWN	2	10	68	20	88
RUR	5	22	43	30	63

Fig 5.3: Student impression of difficulty (%) of science show in school groups

In comparing the three student groups, the urban group comprised 40 Grade 9 students from a former model-C school in a suburb of Richards Bay. There were 28 girls and 12 boys (from 3.4.3). As will be seen from the various graphs in the detailed multiple-choice results (appendix i), students were not very excited about science classes at school, but most of the group (80%) would take Science in Grade 10, despite this lack of enjoyment. The group was polite and well-behaved on the whole, and certainly the most animated of the three groups. Very few of these students had been to the Science Centre before (only 28%), despite their school being just around the corner from the Centre. They all found the show to be either enjoyable (25%) or very enjoyable (75%) and most found it easy (50%) or very easy (48%) to understand. This finding raised the question of whether this obvious confidence would be reflected in their good performance on the various test instruments, which is reported in the next section. They reported that most of the show was new to them which probably added to their enjoyment. Urban students showed greater confidence both in their manner and attitude on the day, which would appear to be a product of their more privileged schooling (see 3.1.2). They seemed unafraid to attempt the drawings, with only 5 students out of 40 on average leaving the drawings blank for each question. Their drawings generally were at least attempts at an explanation and tied in better with their written answers. Across all questions they stated that 60 % of their knowledge had come from the show, compared with 40 % which they had known prior to the show.

The township group comprised 40 Grade 9 students from a modern, well-equipped school in Esikhawini Township about 15 km from Richards Bay. There were 20 girls and 20 boys, all isiZulu speaking students, with English as second language. Again, from the various graphs in the detailed multiple-choice results (appendix i), these students stated that they enjoyed science classes at school (77% finding them quite or very enjoyable), and almost the entire group (90%) said they intended taking Science in Grade 10. The group was interested and well-behaved on the whole. About half of these students had been to the Science Centre before (55%), and they all found the show to be either quite enjoyable (20%) or very enjoyable (78%) and most found it easy (68%) or very easy (20%) to understand. The township students appeared to be moderately confident both in their manner and attitude on the day of the tests. Again this finding raised the question of whether this confidence would be reflected in their performance on the various test instruments, which is reported in the next section. They reported that most of the show was new to them which probably added to their enjoyment. They were apparently nervous to attempt the drawings, with 12 students (30 % of the group) on average leaving the drawings blank for each question. Across all questions they stated that 72 % of their knowledge had come from the show, compared with just 28 % from prior knowledge – the lowest figure of the three schools.

The rural group comprised 37 Grade 9 students from a former Department of Education and Culture school in a rural area not far from Empangeni. All students were isiZulu speakers. There were 19 girls and 18 boys. Half of the students expressed that science classes at school were very enjoyable, and most of the rest quite enjoyable, which is surprising given their lack of resources compared with the others. Most of the group (86%) would take Science in Grade 10. The group was quiet and reserved on the whole, and the least animated of the three groups. Most of these students had been to the Science Centre before (78%), despite their school being some way from the Centre and there being no easily available public transport. They all found the show to be either enjoyable (24%) or very enjoyable (73%) but on the whole fewer found it easy (43%) or very easy (30%) to understand. The rural students showed little confidence in their manner and attitude on the day and it would be interesting to see if this is reflected when their results are considered. As will be discussed below, they were very unwilling to attempt the drawings, with 19 out of 37 students on average (over half of the group) leaving the drawings blank for *each* question. This was unexpected and will need to be more closely monitored in future studies to try to ensure that all students at least attempt the drawings. Across all questions they stated that 60 % of their knowledge had come from the show, compared with 40 % which they had known prior to the show

5.2 Students' conceptions and mental models of sound

This section will be framed by the 4-level framework of Grayson, Anderson and Crossley (2001) (as in table 2.2 in section 2.6) to organise the student results in terms of evidence of such student difficulties. These will be compared for the different school groups and examples given of where they were apparent. These difficulties are ordered and numbered in relation to table 2.2 in and will be used to identify which difficulties were most prevalent in each group. I did not comment on difficulties in the table for which little or no evidence was found from student answers. It was hoped that student data would not only reveal difficulties encountered by the students, but also clear differences between the three student groups. This section will present data from all the probes (OES and MCQ) which were validated in Chapter 4. Section 4.2.3 gives examples of both sound and unsound written answers and further examples can be found in appendix j, which also shows examples of both sound and unsound drawings. Further examples of drawings (both sound and unsound) follow in this section.

Sound travels as a transverse sine wave (Difficulty d1 in table 2.2 in section 2.6)

Urban students shared this conception with all the other groups which was especially evident in the Open-ended Survey (OES) questions 1, 2 and 3 (appendix c). The vast majority of students drew transverse sine waves in response to all three of these questions. There was strong evidence that they had confused the representation (graph of amplitude versus time) with the reality (longitudinal sound wave) as evidenced by their drawings, a few of which even included axes, like figure 5.4 below.

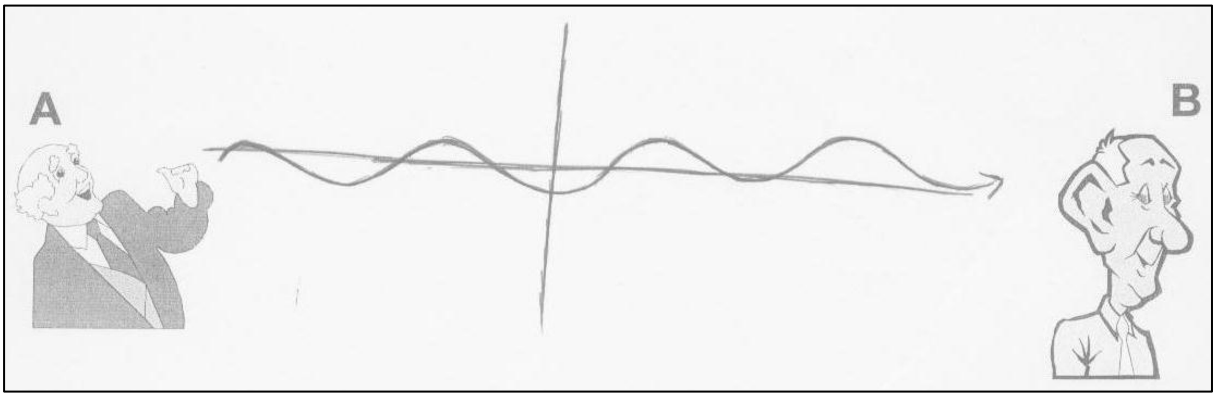


Fig 5.4: Student drawing in response to OES 1 including axes

These findings correspond with those of Wittman (2003), Eshach (2006) and Linder (1992), who similarly found that students misrepresented sound waves as being transverse. In his list of Common Misconceptions about sound, Hapkiewicz (1992) identified one of them as: "Sound waves are transverse waves (like water and light waves)." As these students had not studied sound before, it would appear that this confusion may have been created by the show itself during various activities which were aimed at demonstrating to students the concept of sound traveling in waves (appendix a) Urban students were quite capable of explaining *how* sound waves traveled from A to B in words, with 25 out of 40 giving answers like the following quotation, which were at least partially correct, but this was not reflected in their drawings, like Fig. 5.4 above. This drawing clearly shows a transverse "sound" wave travel from A to B, even with axes. The student was not particular where the sound originated from (it starts closer to A's hand than his mouth) but it does seem to be "aimed at" B's ear.

> (U) "because of sound waves that are produced and therefore travel towards the ear and are received by the ear drum and the sound is registered to the brain for understanding"

This difficulty did not come out clearly from MCQ 2 though, where only 25% of students selected that "Sound travels in transverse waves, despite it being so prevalent in their drawings. At this stage it was possible that naïve students did not associate the word “transverse” with a “sine-wave” drawing, again not making translation between different questions or within the ER.

Township students showed the highest incidence for this conception, with 33 out of 40 students drawing some sort of transverse sine-wave in the Open-ended Survey Question 1 (OES1). What was interesting is that 12 township students drew sharply pointed triangular waves as in Fig. 5.5 below, something I hadn’t encountered before. It would be interesting to find out why they developed that picture (which came up in both OES 2 and 3 as well.)

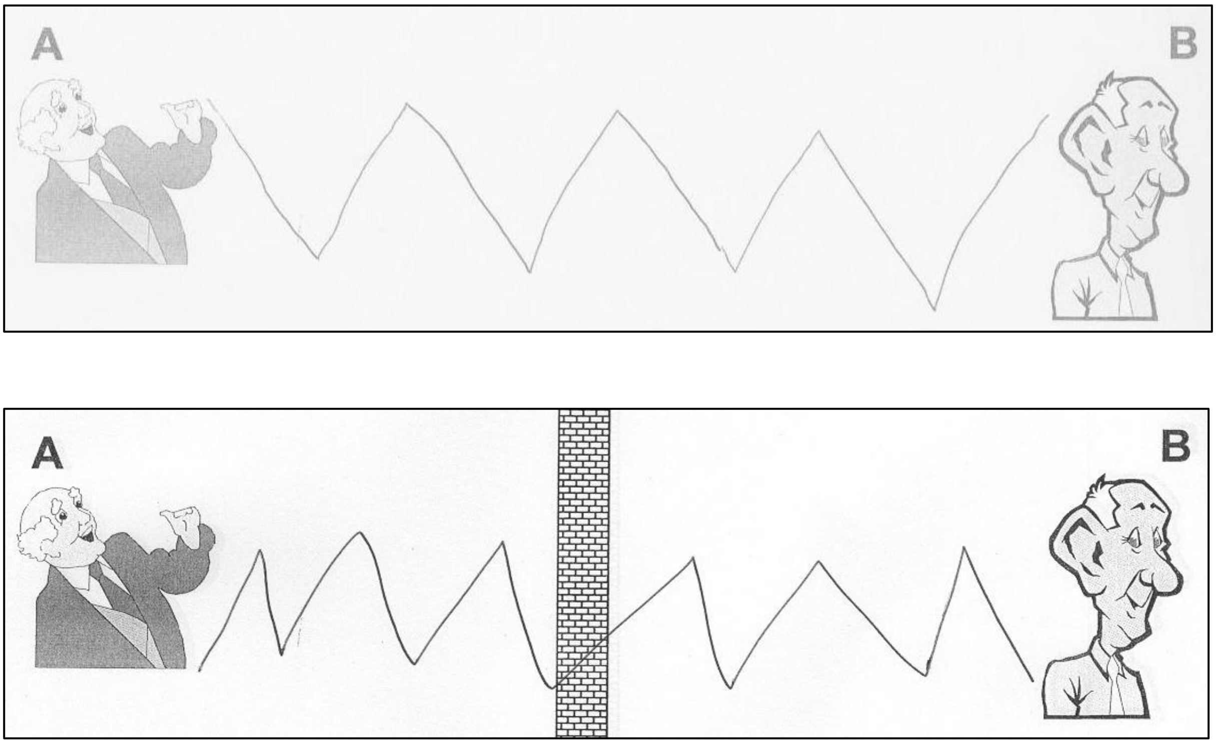


Fig 5.5: Student drawings in response to OES 1 (above) and OES2 (below) with triangular waves

In (OES) 1, 2 and 3 most students drew transverse sine waves in response to all three of these questions. Those who did this may well have learned it from the show itself, as many of the activities displayed sine waves on the screen which seems to have misled the students, as discussed further below. This problem was also revealed in the verbal student responses and in the MCQ with 35 % in MCQ 2 choosing that sound travels are transverse wave. A similar percentage was recorded for the pre- and post-tests. Written explanations were weak and students clearly used words they didn’t fully understand, such as the following:

> (T) “A have to use frequency (Pitch) so that B can hear what he is saying”

To some extent, **rural** students shared this conception with all the other groups, but it was far less pronounced in their case. In the Open-ended Survey (OES), questions 1, 2 and 3, some students drew transverse sine waves in response to all three of these questions but there were many other representations as well. The rural group tended to abstract less and focus more on concrete explanations, so this problem of confusing the representation (graph of amplitude versus time) with the reality (longitudinal sound wave) was less evident. Those who did this may well have learned it from the show itself, as detailed in below. It also came out verbally with 24% in MCQ 2 choosing the option that sound travels are transverse wave. Finally, four rural students simply redrew the people in the picture, without adding any waves, indicating that they had not understood the instruction.

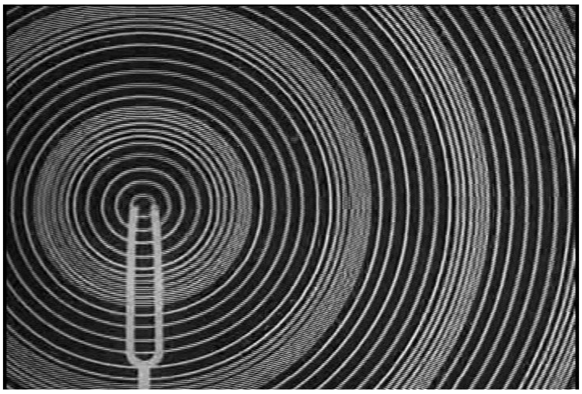
It might have been useful to further explore rural answers from their written explanations to get more input towards the R-M factor, but these were so poor that they were mostly unusable. I only judged three of their answers to be even partly correct and even their best answer focused mainly on the overt features of the sound i.e. how loud it is, rather than on how the waves travel, as shown by the following typical student response:

> (R) "A is shouting at B and the space between them is not that big but I am sure that B can hear A, cos sound can travel in gases and it is a wave that we cannot see but I saw it on a laptop. And it can also depend how loud is A is speaking to B"

Only 3 of 37 responses mentioned waves at all and showed great confusion between terms, such as the following:

> (R) "sound travels using waves from A to B or vibration which makes the sound higher"

As this particular difficulty (d1) was so prevalent for all three groups, and evident in OES 1, 2 and 3, it warrants further discussion: urban and township students were fairly united in producing the *representation* rather than the *reality*, but rural students (who attempted the drawing) came up with many different pictures, some unique, including simply redrawing the people from the picture I supplied or drawing a pipe which carries the sound. All students thus showed that they lacked the R-M factor of ability to reason with the ER itself and to translate between the different visualizations in the ER. The overall picture was in sharp contrast to the picture I would have liked to see of circular wave fronts traveling like ripples in a pond (which no student produced here, or in either of the pilot studies). These pictures and photos are common in textbooks and general publications, like for example the two below in Fig 5.6 obtained off the internet.



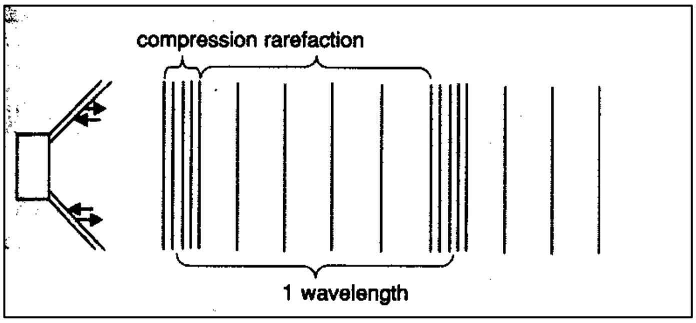


Fig 5.6: Depictions of sound traveling in longitudinal waves (from Google Image)

The closest that a student came to this was in this drawing of parallel wave fronts (figure 5.7)

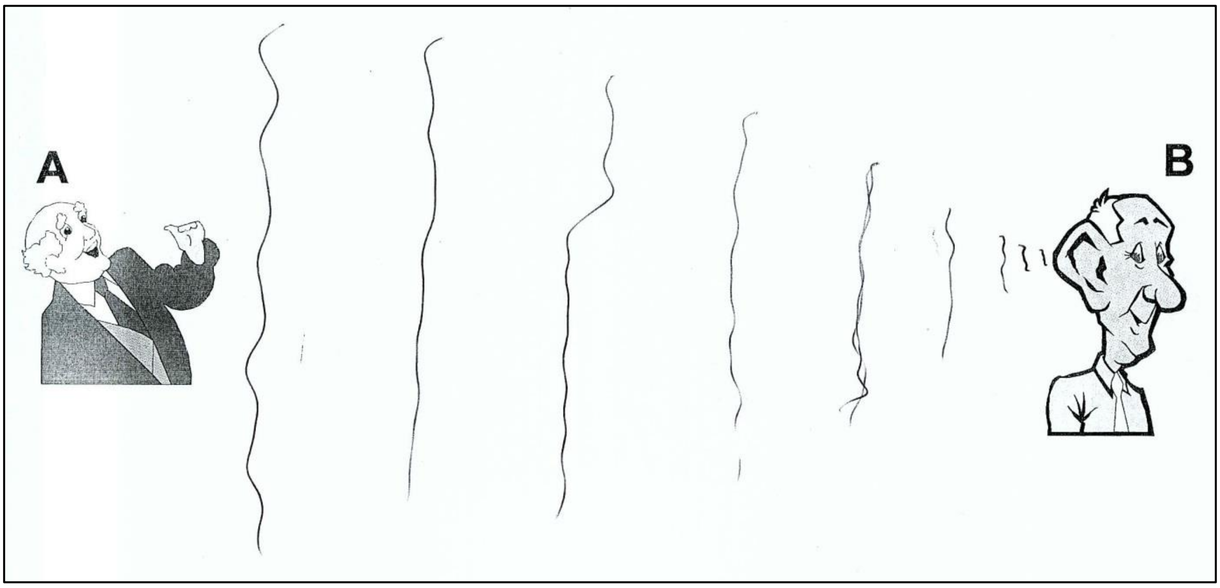


Fig 5.7: Student drawing in answer to OES 1

Considering that these students were Grade 9's, who have not yet studied sound and waves in school science, it was not surprising that they had not seen these correct pictures. Although, Linder (1992) identified that textbook diagrams are often misleading or even false, it still came as a surprise that I had probably caused the development of this misconception at various times in the show. Looking more carefully at the PowerPoint slides in the show, I realized the reason for this problem was that eight of the slides showed sound waves as transverse rather than longitudinal, and there were no slides showing sound propagating via longitudinal waves. More students (52) identified their source of knowledge as coming from the show rather than before the show (38 students), and their answers seemed to reflect what they had learned from the show. These slides are shown overleaf in Fig. 5.8.

In retrospect, this problem should have been picked up at the Pilot stage and the presentation altered or the problem addressed. In the initial pilot, *all* of the students drew transverse waves of some sort in response to the drawing requested in question 7 – see appendix g. A summary of their drawings for question 7 can be found at the end of appendix m. Even in the open-ended survey pilot, student responses to the drawing questions invariably showed transverse, rather than longitudinal waves. In both cases, my focus was on the probes themselves, rather than the show. In the first case, I was using the probes to check for discernible differences between the three student groups, and in the second I was refining the probes and looking for ambiguities or unclear questions.

So it seems clear that the very nature of the ER (M) is causing this main student difficulty. Later in this study I shall use student answers to shed light on learning during the show in order to improve the show.

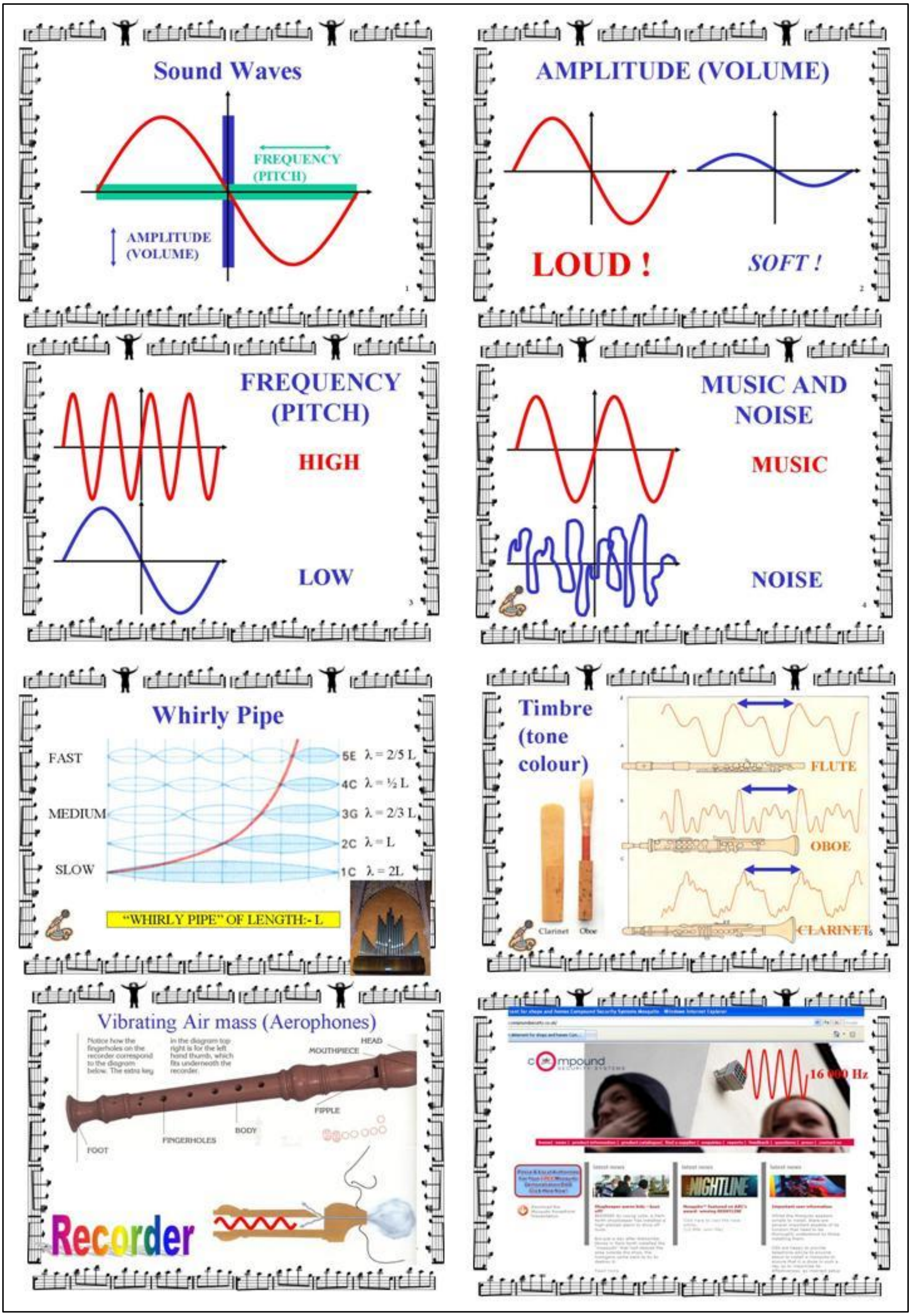


Fig 5.8: Eight PowerPoint slides from the science show depicting transverse waves

While the top 6 slides in Fig. 5.8 are attempting to show a *representation* of a sound wave (rather than a *real* sound wave), it was not possible for the students to know that, and they understandably believed they were looking at real sound-waves in their literal interpretation of the diagrams. The bottom two slides in Fig. 5.8 actually show transverse sine-waves emanating from a recorder and a loudspeaker, further confusing the point, and clearly needing to be corrected.

A large part of the show (see appendix a) was concerned with using a computer-based oscilloscope to allow students to *see* sound waves (using the “Data Studio” software from PASCO Scientific). But what they saw here was a graphical representation using transverse sine-waves to present concepts like frequency and amplitude and to show the difference between music and noise. Typical screens are shown in Fig. 5.9 below:

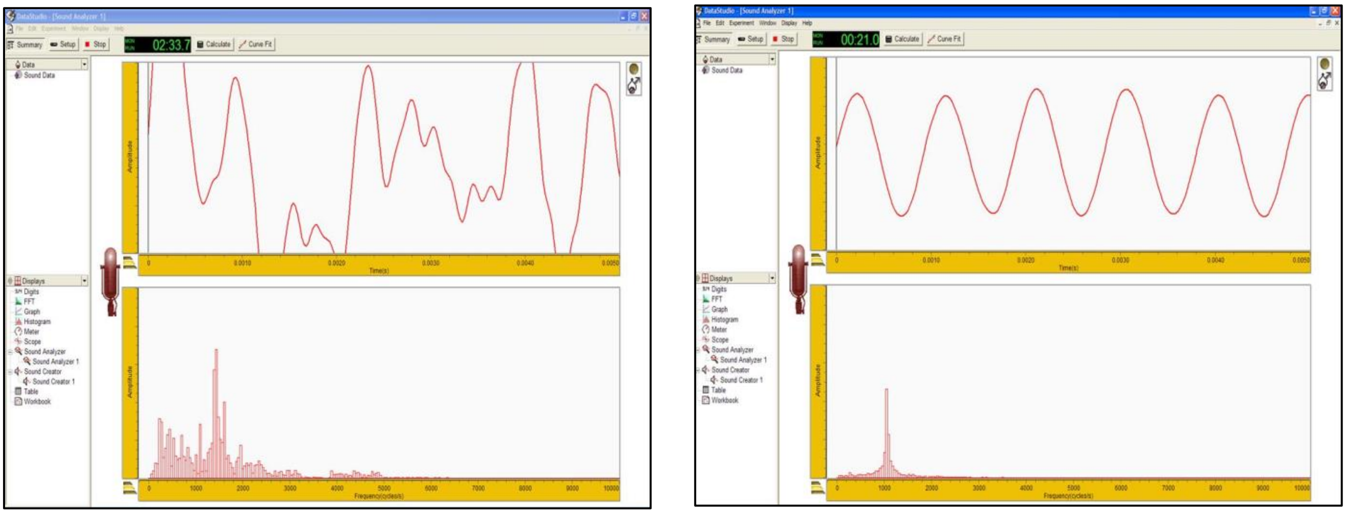


Fig 5.9: Typical “Data Studio” screens seen by students in the science show.

Although the two screen shots shown in Fig. 5.9 are in fact graphs of amplitude against time, representations of a sound wave, the students were clearly not aware of that and faithfully reproduced sine waves, often even supplying the axes, and at least one student from each school produced a picture very similar to the screen shot above – even including something which looked like the frequency spectrum at the bottom (Fig. 5.10).

Fig 5.10: Student drawing in response to OES 1

This confusion between the reality and the representation (Linder, 1992) is common in science education, and also causes confusion with motion graphs, vector diagrams and other representations. As can be seen in the next two questions (OES 2 and 3), the “sine-wave-sound wave” was repeatedly used, and its limitations were shown when sounds needed to turn a corner or pass through a wall. The show should certainly include a diagram early on of sound as a longitudinal wave, and make clear that sine-wave pictures which follow are representations or graphs, and are not a snapshot of the actual wave traveling. Even better would be some demonstrations (perhaps with a slinky-spring) of longitudinal waves.

This will have to be carefully done, as even the reality plus representation pictures common in textbooks like Fig. 5.11 below are very difficult for students to interpret. Even to an expert, it is not clear *what* is being graphed in 5ig. 5.11: (pressure, velocity or displacement?) and against what (time or horizontal distance), no axes are supplied and the two are difficult to match up, even for teachers. Furthermore the lower graph is representing displacement from the equilibrium position *vertically*, at 90 degrees to how it is in the longitudinal wave above it, which is very confusing.

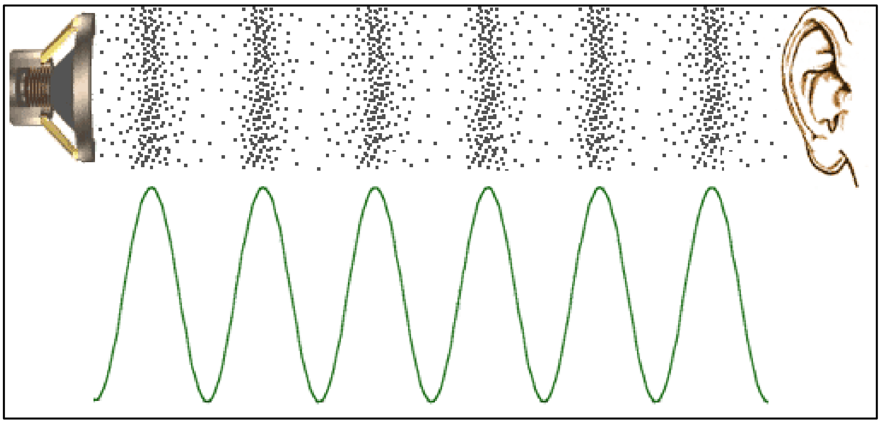


Fig 5.11: Illustration of sound propagation showing both reality and representation (from Google Image)

The other problem with students drawing these sine-wave representations is that practical considerations were often ignored. Waves did not necessarily start at A's mouth nor end at B's ear, or they were "disconnected" in the middle. Township students (almost) uniquely drew sharply pointed triangular waves and it would be interesting to find out why they developed that picture. Other drawings (for example fig 5.12 and 5.13 below) also included some of the sound "threads" (Fig. 5.12) and "bubbles" (Fig. 5.13) identified by Eshach (2006) (and perhaps learnt from cartoon strips). In all of the drawings (about 100) I found no evidence of the "reverse sound" (i.e. from hearer to source) which was reported by Boyes and Stannistreet (1991).

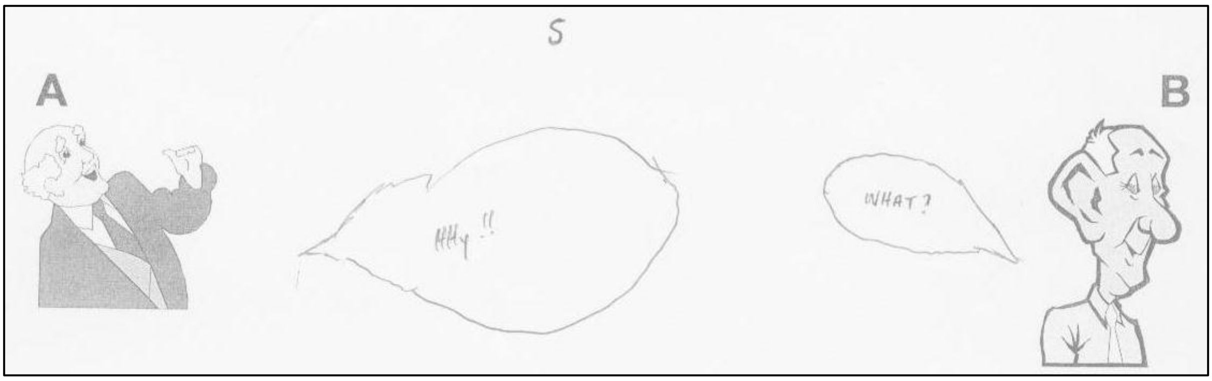


Figure 5.12 Student drawing in response to OES 1 showing "sound bubbles"

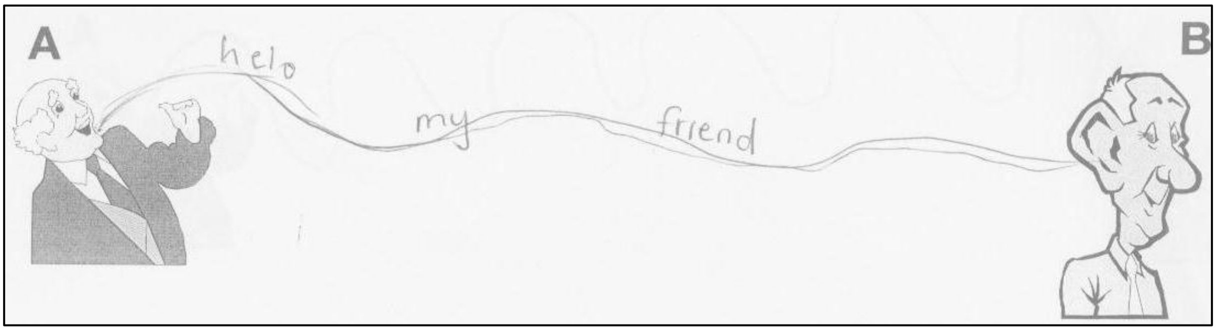


Fig 5.13 Student drawing in response to OES 1 showing "sound threads"

I was interested to look for examples of the different conceptual models which students create for sound, especially as summarised by Hrepic et al. (2010), who suggested that students begin with an "entity model" (Linder and Erickson, 1989), and move towards a wave model as their scientific knowledge progresses. In between they will often form "blend models" which join some of the ideas from both of the other models. What is noticeable from the previous descriptions in school types is that urban students describe sound far more in wave terms, township students less so, while rural students barely use the word "wave" at all. Township students occasionally seemed to indicate some sort of an entity model for sound such as in the following quotation:

(T) "sound is a regular shape"

but it was not clearly discernible. Most of Linder's participants were interviewed, allowing him to probe with follow up questions and to get to the heart of their conception of sound. Remembering the comment of Hrepic (2010), that: **"**… students rarely displayed coherent reasoning at the level of a structured model without being prompted with additional questions" I would argue that a single written answer is not enough to bring out a complete model of sound from a naïve student.

There were also some unexpected nuances. Township and rural students both commented on the fact that the men in the picture (or at least the hearer) were *old* men, so would struggle to hear. This is illustrated in the following quotes and probably was due to the propensity of these students to look for naturalistic explanations in keeping with their life experiences.

(T) "my friend I think doesn't like this noise it sound so un bearable and person B is older"

(R) "if A can talk at a low but not very low voice will hear him because he is old and can only hear sound that have a low, but not very low frequency"

While a few students mentioned air, none talked about particles or molecules, so none got close to Linder and Erickson's (1989) microscopic entity model.

Sound needs holes to escape a container (d5)

This conception was evident from written answers to OES 2, as **urban** students mentioned sound traveling through the spaces between the bricks. (It was not easy to decide from the *drawings* if this conception was present, but the drawings did lead me to establish a new student conception – see d18 later in this section)

> (U)"The sound is able to travel through a wall due to the bricks having spaces in between them and the air is present there, and thus making B hear due to "imperfections" in the wall."

Many students, though, correctly spoke about the sound vibrating through the wall:

> (U) "The Sound travels from person A's mouth and it vibrates through the wall, transmitting sound to person B."

This conception was only evident from 3 out of 40 **township** students, so not very common. An example is the following quote:

> (T)"The sound is going to travel through small spaces of the bricks where two bricks have been meet "

Four township students came up with the notion that air is everywhere and therefore sound can get everywhere: this notion that "sound is air" was found by Caleon and Subramaniam (2010) and relates to d9 (table 2.2). It is interesting that these same four students also answered question 9 correctly, stating, as apparent in the following quotes, that you cannot hear sounds in space because sound needs air to travel through, at least showing consistency.

> (T) "They can hear each other from around the corner because sound can even travel in air and air is everywhere."

> (T) "Sound is like air you can hear it even though you are in bedroom"

Rural students did not show clear evidence of this difficulty, but again their extremely weak written answers made it difficult to tell. In this case the problem of sound travelling through a wall was not directly addressed during the show and so students' difficulties are more an indication of their prior knowledge and mental models (C) than what they learned from the show. There were a few drawings here showing parallel wave fronts interacting with the wall and diminishing in size after going through it, like fig. 5.14 below.

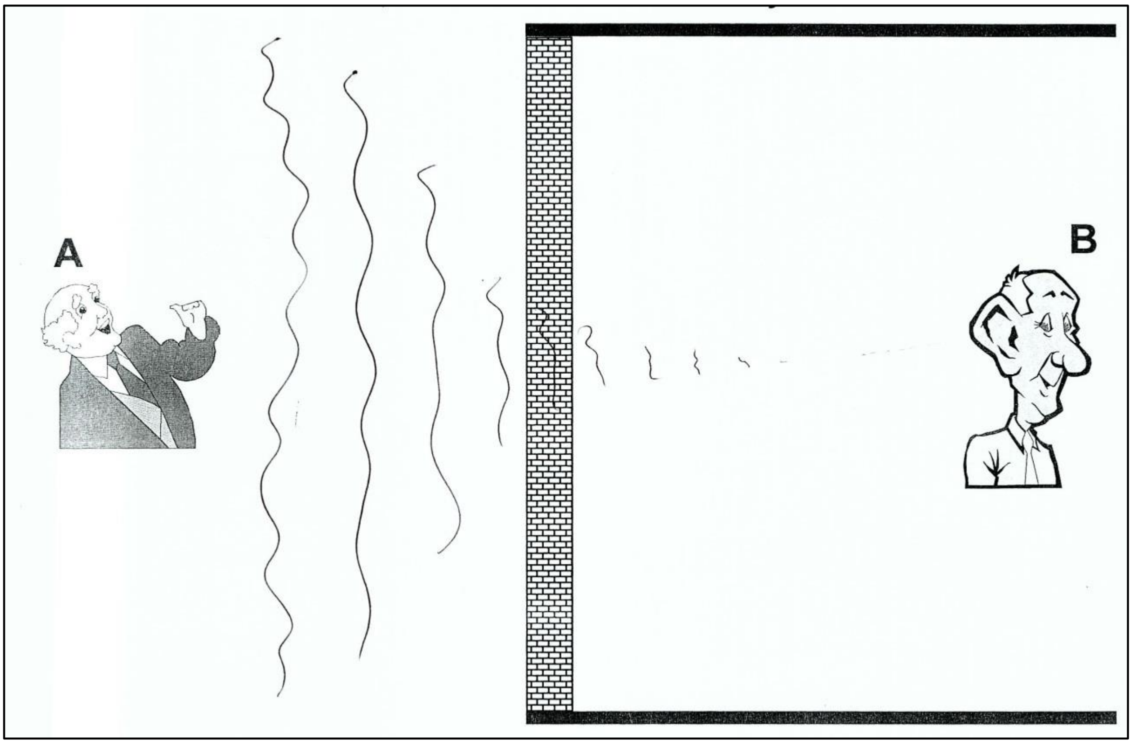


Fig 5.14: Student drawing in response to OES 2

Sound can travel through a vacuum and therefore through space (d6):

The OES 9 probe revealed that quite a few (14) **urban** students believed that sound travels like light (as an electromagnetic wave (d 17)) and can propagate through space., as this written answer attests:

> (U)"Sound is part of the electromagnetic spectrum and can move anywhere (radio waves)."

This relates to MCQ 2, where 72% chose that "sound travels in Electromagnetic Waves" (d 17). In the post-test, students almost completely abandoned their image of sound travelling as an electromagnetic wave (from 72% to 12%) for the correct longitudinal wave (from 2% to 72%).

Again, there were a good number of conceptually sound answers such as the following:

(U)" In outer space there is no oxygen (air) in that way sound cannot travel because sound needs air in order for it to travel."

Again here the problem of sound travelling through space was not directly addressed during the show and so students' difficulties are probably an indication of their prior knowledge and mental models (C) rather than what they learned from the show. Drawings from students were mostly illustrations rather than explanations, like the one in fig. 5.15 below.

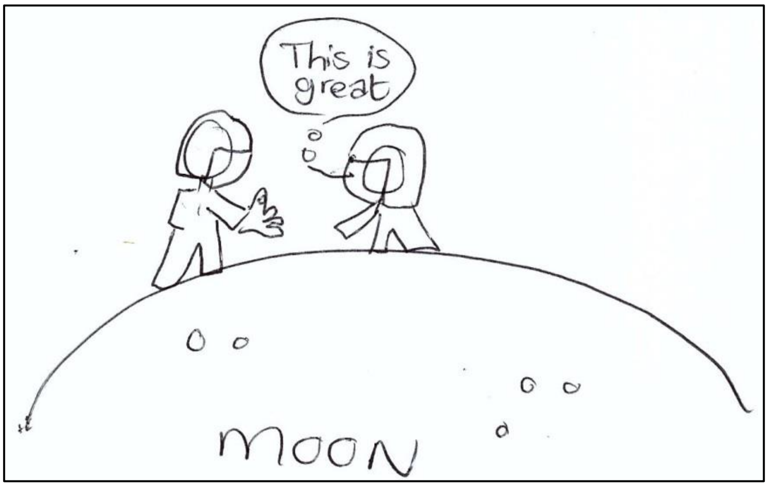


Fig 5.15: Student drawing in response to OES 9

A few students did produce nice drawings indicating that they understood the need for air in order for sound to travel.

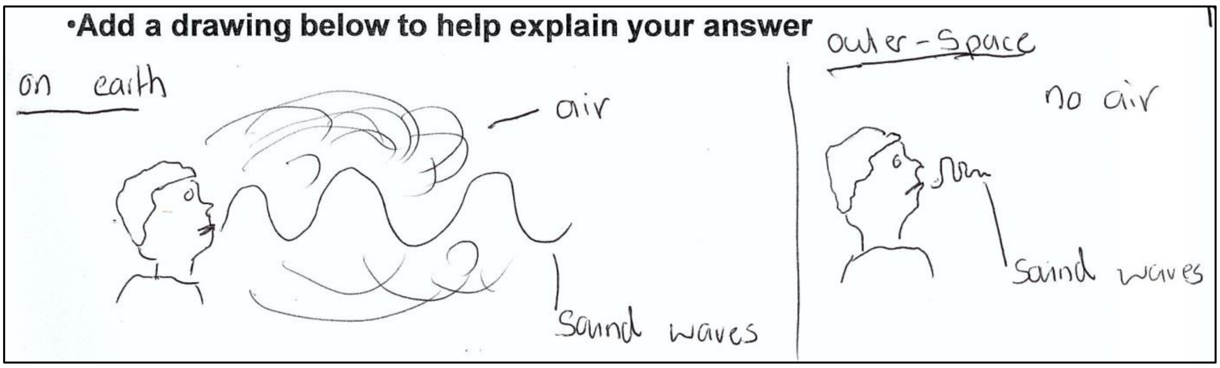


Fig 5.16: Student drawing in response to OES 9

Confusion between volume and pitch (and amplitude and frequency) (d8):

This difficulty was revealed in MCQ 6 and 7 among the township and rural students (especially in their pre-tests – C), but there was no apparent student conception of this nature evident among the **urban group. Their** answers were spread quite evenly over all the options. For example, urban students mentioned in OES 7 that loud sounds move the eardrum too fast (6) and that the frequency is too high (3) and it affects our ears (2). Despite comparatively well-written answers, the same confusion was apparent with the concepts of loudness and high frequency being considered synonymous. This confusion was also noted by Hapkiewicz (1992) and Caleon and Subramaniam (2010). This confusion between different wave properties was evident in several answers, like the following triple confusion in answer to OES 1 (This answer confuses pitch, wavelength and volume as being the same thing):

(U) "The higher the pitch (wavelength) the louder the sound is"

In MCQ 6: 52 % of **township** students answered that "the wave property related to volume is frequency" (rather than "amplitude" only 12%), possibly showing a confusion between volume and pitch (and amplitude and frequency). This dropped to just 28% in the post-test, while the correct answer rose to 60%, indicating that this conception was mostly dealt with in the show and showing good ability to integrate new knowledge learnt from the ER (R-C)

Again this confusion arose in MCQ 7 where 52 % of township students answered that "the wave property related to pitch is wave-speed" (rather than "frequency" only 15%). This dropped to just 22% in the post-test, while the correct answer rose to 50%, indicating that this conception was partly dealt with in the show This question shows some evidence of the same confusion as in question 6 above, but students did choose "wave-speed" indicating perhaps a belief that high-pitched sounds travel faster. This was also found in the rural group but not in the literature. This confusion is perhaps a difficulty in distinguishing relationships between quantities (volume and pitch) in the ER and needs to be addressed.

Drawings for OES 7 were not very good, but a few did indicate that students understood that it was the amplitude of the wave ("large vibrations") which caused pain to the ear (fig. 5.17).

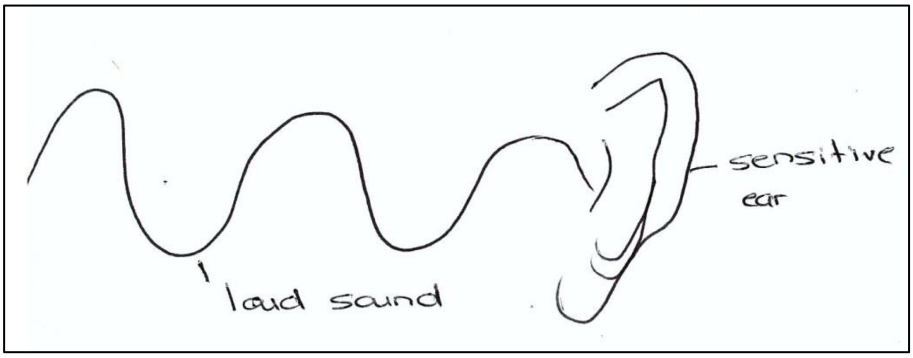


Fig 5.17: Student drawing in response to OES 7

Sound can only travel if air is present (and therefore can't travel through liquids or solids) (d9)

OES 5 was badly answered by **urban** students, most of whom believed that sound travels fastest in gases (air in particular). This also links in with d4 that the speed of sound in a medium is linked to the "sound-resistance" of that medium. In their written explanations, these students stated that the particles of a gas are far apart (10) or that gases are light or less dense (5) or their particles loosely bonded (2). Some (5 students) said that a gas doesn't offer any resistance (also disruption, obstruction, interference) to the sound with one even mentioning that solids offer a "sound resistance", echoing the terminology of Linder (1992) in his study on student conception of the speed of sound. A few students mentioned that solids and liquids are heavy and that their particles are strongly bonded and have no spaces between them, such as in the following quotation:

> (U) " Solids can either block off sounds or make it very soft, liquids often produce a sound of their own and this can also be a form of sound resistance. Also liquid atoms are relatively close together and not much can move through the gaps much like a solid. In a gas, however, there is far more space between the gas particles for sound to travel through. Gas does not form any resistance against the sound."

MCQ 3 also revealed these urban student conceptions related to the speed of sound, with 88 % of students stating that sound travels fastest through air, rather than steel (which no-one chose) or water. The misconception that sound travels fastest in air was *not* corrected by the show as expected since this was deliberately left out of the show as a comparison question. Again as this was not directly addressed during the show students' difficulties are more an indication of their prior knowledge and mental models (C) than what they learned from the show. Different notions on what affects the speed of sound in a medium were studied extensively by Linder (1993) and is item 4 on my difficulties table. It possibly also echoes the conception found by Caleon and Subramaniam (2010) that there must be air present for sound to travel (item 9 on my difficulties table). Very few students indicated that “sound travels fastest through a vacuum” (10%), item 6 on my difficulties table.

OES 5 was also poorly answered by **township** students (although better than the two other groups), most of whom believed that sound travels fastest in gases (air in particular). This could also link in with d4 that the speed of sound in a medium is linked to the “sound-resistance” of that medium. Township students did not really use a particulate model in their drawings (most leaving it blank) or their explanations. Most stated here that a gas is like air, and sound travels fast through air. Again, object based explanations were preferred to wave-based.

MCQ 3 also revealed these township student conceptions related to the speed of sound, with 68 % of students stating that sound travels fastest through air, rather than steel (only 1 student) or water. The misconception that sound travels fastest in air was again *not* corrected by the show:- in fact this choice rose to 78% in the post-test. A few students indicated (a) that “sound travels fastest through a vacuum” (20%), item 11 on my difficulties table.

OES 5 was poorly answered by **rural** students; most of whom believed that sound travels fastest in gases (11) or liquids (13) than in solids (6). Their drawings did not show a particulate model of matter and usually showed actual instruments or objects which were solid, liquid or gas. An example of such a drawing is shown in Fig 5.18 below. As can be seen, the drawing did not display a particulate model of matter and reflected more the student’s everyday experience than a scientific model.

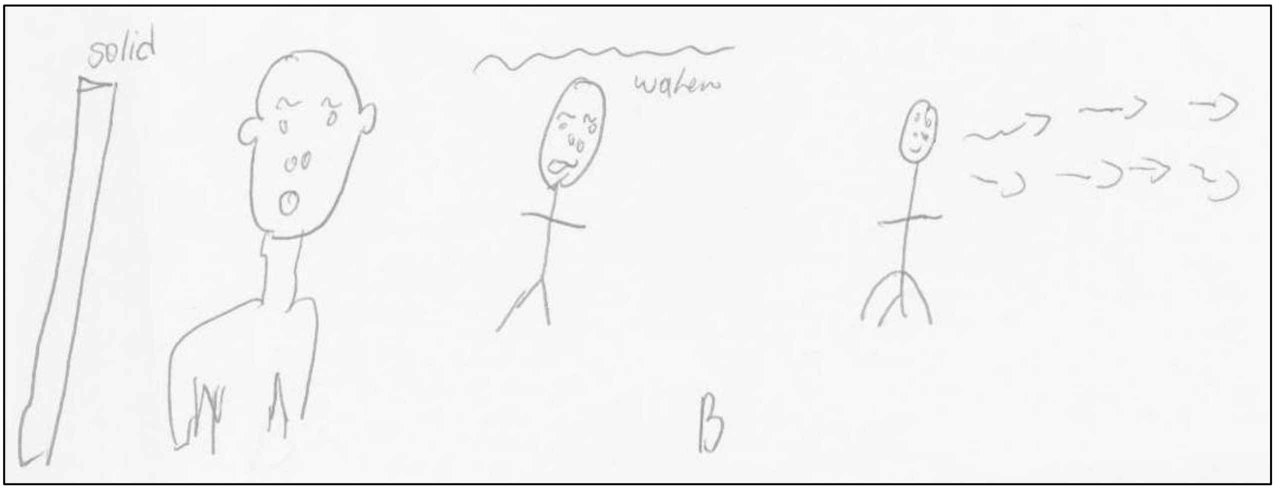


Fig 5.18: Student drawing in response to OES 5 showing instruments and people

A few drawings did display particles (mostly from urban students) but often these produced a "reverse" answer, where the phases with more space between the particles allowed the sound to travel faster (fig 5.19).

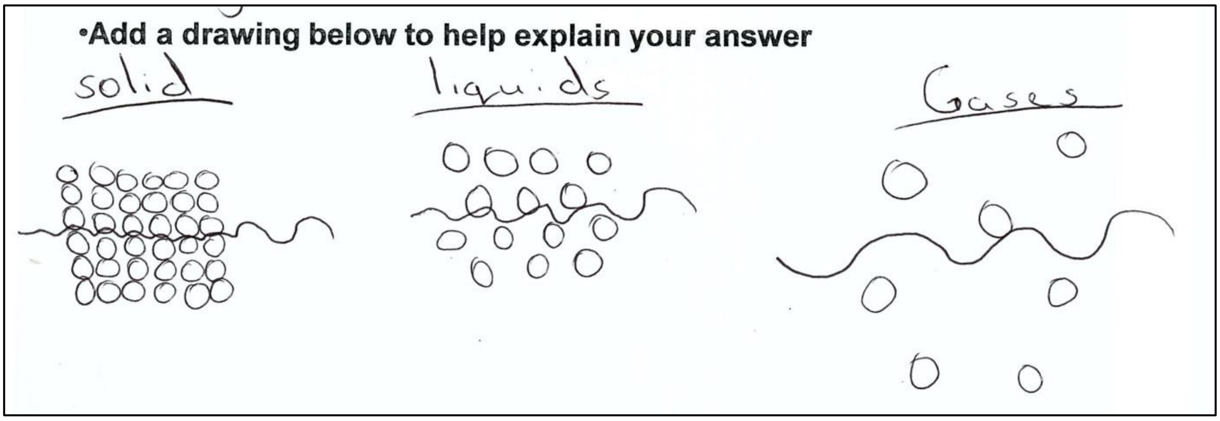


Fig 5.19: Student drawing in response to OES 5

This also links in with d4 that the speed of sound in a medium is linked to the "sound-resistance" of that medium. Sadly their written explanations were so poor as to add little to the subject.

It was also interesting that in OES 5 the option: "Sound travels fastest in liquids:" was mainly chosen by rural students and was their most popular choice (13 students). Unfortunately, their language made it very difficult to make sense of their answers to discover why. There was no real agreement between any of the answers offered, but some of the reasons given were that liquids flow very easily or quickly, are soft or are faster than solids. One student added that rain falls fast. These would make interesting subject for further investigation.

MCQ 3 also revealed these rural student conceptions related to the speed of sound, with 73 % of students stating that sound travels fastest through air, rather than steel or water. The misconception that sound travels fastest in air was again NOT corrected as it was not covered by the show.
Music has low volume and noise has high volume (d13)

This difficulty was not evident in the answers for the **urban** students. As in d8 they didn't seem to have a problem distinguishing between two quantities (music and noise here, volume and pitch in d8). In contrast, the township group did show this misconception. In OES4 the choice of words used for music by **township** students was interesting, with 20 students describing it as smooth 12 as soft. Very few students from other groups used the word "soft" (only 2 urban and 3 rural students) and this would appear to indicate the conception associating music with low volume and noise with high

volume. In addition, 12 township students described noise as being “loud or irritating”.
Answers to MCQ 8 revealed that many **rural** students felt that music is different from noise in that its wave has a smaller amplitude (32%). An almost equal number (35%) attributed it to music having a higher frequency, so again this may be ignorance and confusion more than a clear student conception. In OES4 rural students again spoke in terms of the *effects* of the noise, rather than wave-shapes, with answers like:

> (R) “…noise doesn’t make sense but music make sense and it is smoothly it sound nice and it doesn’t cause you a headache.”

This concept (that music has a smooth, repetitive wave shape) was carefully demonstrated using the “Data Studio” software. The fact that the urban students did not show problems here indicates their better ability to reason within the ER (R-M) – especially with regard to the computer simulations which were more familiar to these students than to the other groups. Thus assimilation appears to have been easier for the urban students as the use of graphical features in the ER (M) was more familiar. Many students drew nice representations of the smooth wave associated with music and the rough wave associated with noise here (fig. 5.20)

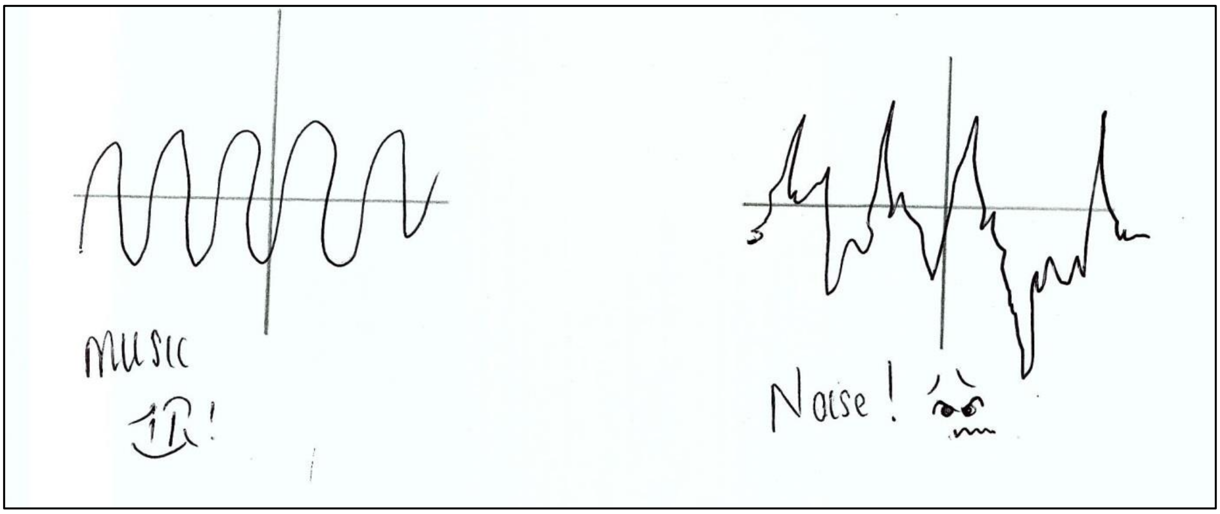


Fig 5.20: Student drawing in response to OES 4

Longer objects vibrate faster, or produce higher notes (d14)

This difficulty was very evident in MCQ 9 in the pre-test, especially the notion that longer objects produce higher notes, with options (a) and (c) being most popular choices for all school groups. All groups improved in the post-test, with the difficulty being mostly sorted out for urban and township schools, but not completely for rural schools, with the most popular choice still being (c). As this was carefully taught in the show using a number of different instruments, this difficulty is perhaps more with the language of the question than the nature of the ER in this case. This particular MCQ is discussed in more detail at the end of the next section (5.3)

The sound box on an instrument is to make the sound clearer or higher (d15)

Answers to OES 6 showed most **rural** students stating that the box makes the sound higher. While this may be a wrong conception, it is just as likely to be an English language problem, as I noticed (both verbally during the show and in marking written responses) great confusion between “louder” and “higher”, which seemed to be used interchangeably. Again more likely a problem of language than with the ER itself, as urban and township students had few problems here. As stated in 4.2.4 this question needs to be carefully reworded to avoid this confusion and answers like the following quotation:

> (R) “ …make the sound higher so that the sound … heard in our ears properly”

Most of the drawings in response to OES 6 were just pictures of guitars, but some showed a pleasing comparison of waves produced with or without the guitar box (fig. 5.21)

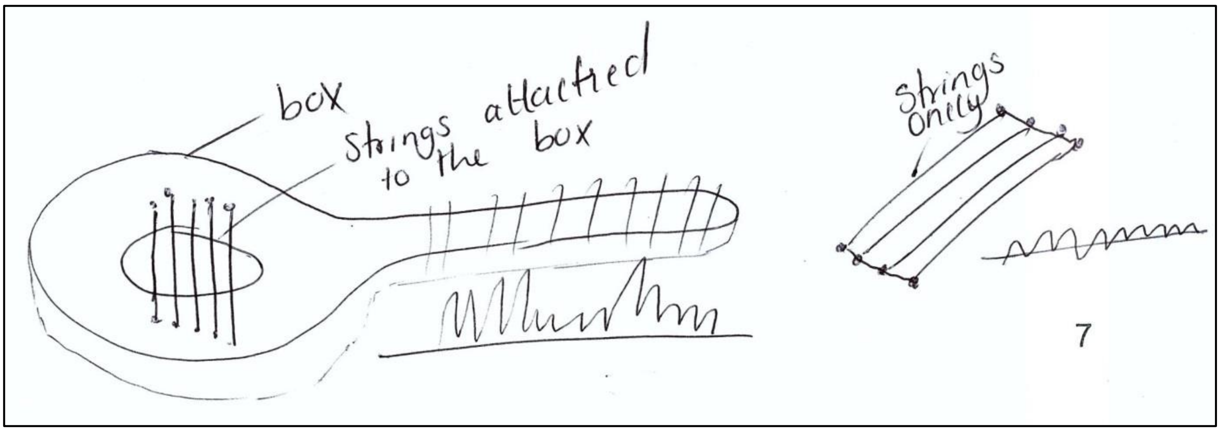


Fig 5.21: Student drawing in response to OES 6

*Vibrations and waves are the same thing (*d16)

In MCQ 1, 65% of **urban** students answered (a) that “a vibration is a wave” (rather than (c) “a regular movement backwards and forwards” only 18%), echoing a confusion also reported by Hrepic et al. (2010) and item 16 on my difficulties table. This dropped to 48% in the post-test, but was still their most popular choice, so this conception was not fully dealt with. Although vibrations were the basis of the show (“Good Vibrations”) the difference between a vibration and a wave was obviously not made fully clear to them. This is an example of where the ER did not satisfactorily explain this concept so that their prior knowledge (C) could be addressed and new knowledge assimilated (R-C). In MCQ 1: 48 % of **township** students also answered (a) that “a vibration is a wave” (rather than (c) “a regular movement backwards and forwards” only 20%), This dropped to just 22 % in the post-test, while the correct answer rose to 65%, indicating that this conception was mostly dealt with in the show, so new knowledge was successfully assimilated (R-C). This problem was not evident amongst rural students, who answered MCQ 1 correctly both in the pre- and post-tests.

The following three difficulties (d17 – d19) were not in table 2.2 as they did not arise from the literature or from my prior experience. They were unexpected (level 4) arising from student data.

Sound travels in electro-magnetic waves (d17)

This newly classified difficulty emerged from MCQ 2. Here 72 % of **urban** students answered (b) that “Sound travels … in electromagnetic waves” (rather than (c) “longitudinal waves” only 2%). This was dramatically reversed in the post-test with this choice dropping to just 12 % and the correct choice rising to 72%, evidence that this conception was mostly dealt with. There was slight evidence of this difficulty in the township and rural group but it was not very prevalent. Again an example of incorrect prior knowledge (C) being successfully addressed by the ER (R-C).

Sound is unaffected by obstacles (d18)

This new, unanticipated difficulty was categorized from results to OES2 (and also OES 3). In OES 2, half of the **urban** students visualized the waves as simply passing through the brick wall unaffected, as if the wall was not there. The drawings of the *same* urban student for question 2 and then question 1 are shown in Fig 5.22 below. One can see that the presence of the wall makes no difference to the student’s drawing.

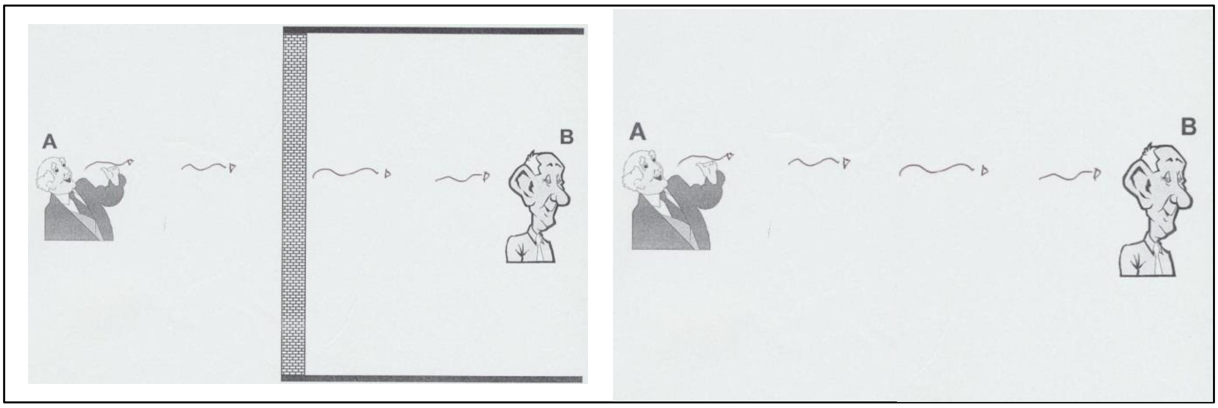


Fig 5.22: Student drawings (from the same student) in response to OES 2 (left) and OES 1 (right)

Some students did show the waves (albeit transverse waves) becoming smaller in amplitude after passing through the wall, hopefully indicating that the amplitude (volume) would be reduced as in Fig. 5.23 below.

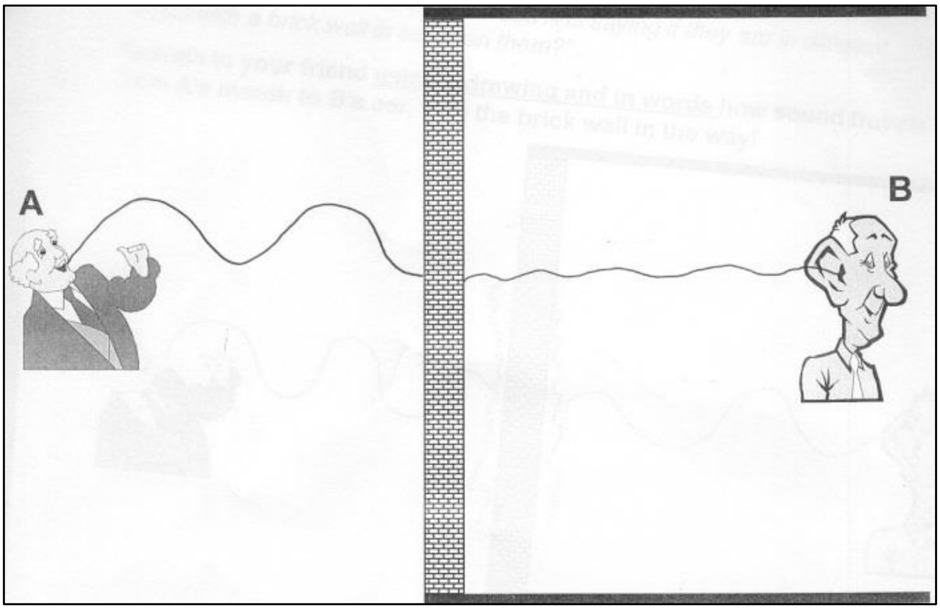


Fig 5.23: Student drawing in response to OES 2 showing reduced amplitude

In OES 2, most (22) township students drew the waves as simply passing through the brick wall unaffected, as if the wall was not there (in a similar manner to urban students). In OES 3, one wrote:

(T) "A corner is just a material but you can still hear through it."

Three ingenious students showed the sound creeping around the open side of the room, perhaps indicating their belief that the wall would affect, or block, the sound, as in Fig. 5.24 below:

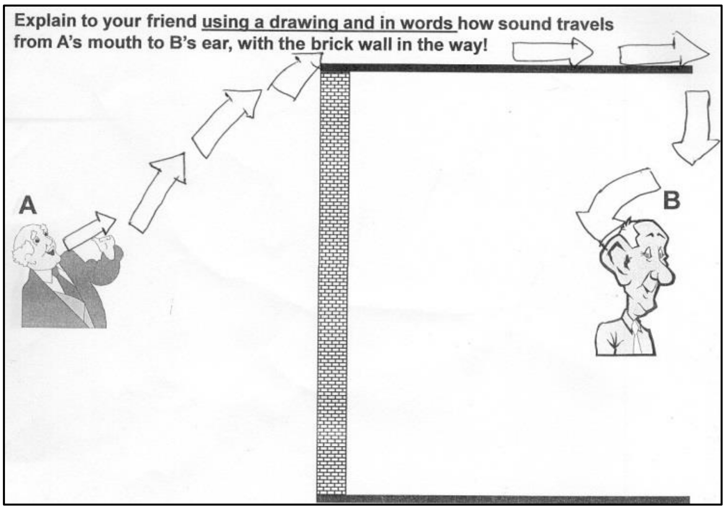


Fig 5.24: Student drawing in response to OES 2 showing sound going around the wall

Only 4 of the 40 responses mentioned waves at all and 6 mentioned vibrations. Again many students mentioned the loudness of the sound and the fact that B would not hear unless A spoke loud enough.

In OES 2, a few **rural** students (6) visualized the waves as simply passing through the brick wall unaffected, as if the wall was not there. This perhaps explains why no students mentioned sound passing through holes in the wall (d5) as perhaps they didn't see the presence of the wall as having any effect at all? Again in OES 3 they allowed the sound merely to travel through the wall instead of having to go around it:

> (R) "It can travel in air the position doesn't change anything the wall is just a shape."

This example relates to d5 and was not addressed directly in the show, so again student difficulties here are a product of their prior knowledge and mental models related to sound (C).

Sound waves will turn to reach a hearer (d19)

This new, unanticipated difficulty was categorized from results to OES 3. In OES 3 **urban** students showed sound waves mysteriously bending at 90 degrees, or curving in order to reach the hearer, with no apparent agent for the turning. This is linked to d6 that sound waves cannot be refracted or bent like light, but urban students didn't show any signs of illustrating refraction, but rather just turned the waves at will. Even their written answers did not reflect any agent for the turning.

(U) “The sound waves travel from Person A’s mouth and around the corner to person B’s ear”

Three urban students mentioned the word “echo” in their answer, but this was not mentioned by any of the other school groups. Their drawings most clearly reflected the “hearer-seeking” sound waves, as in Fig. 5.25 below. One of the problems with the sine-wave depiction is that it causes problems when trying to explain questions like OES3.

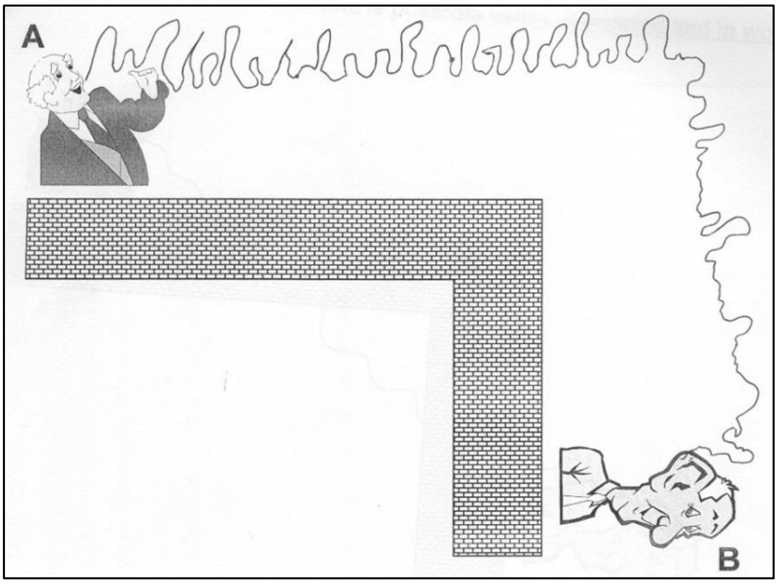


Fig 5.25: Student drawing in response to OES 3 showing sound waves turning towards hearer

One urban student did draw some sort of parallel wave fronts spreading around the corner, but his picture was confused by the direction of his wave fronts (fig. 5.26)

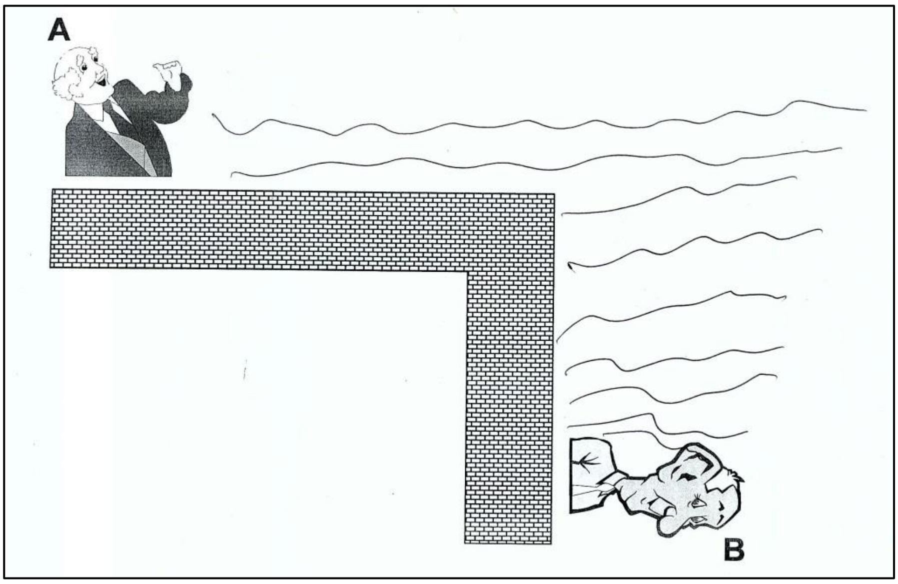


Fig 5.26: Student drawing in response to OES 3

In OES 3 most (23) **township** students also showed sound waves mysteriously bending at 90 degrees, in order to reach the hearer, with no apparent agent for the turning (with very similar drawings to those shown for the urban students above). Even their written answers did not reflect any agent for the turning. Only one student had the idea of the sound wave (albeit transverse) "spreading" around the wall in both his drawing (fig 5.27 below) and written answer.

(T) "carries on with a widen up sound to an open space."

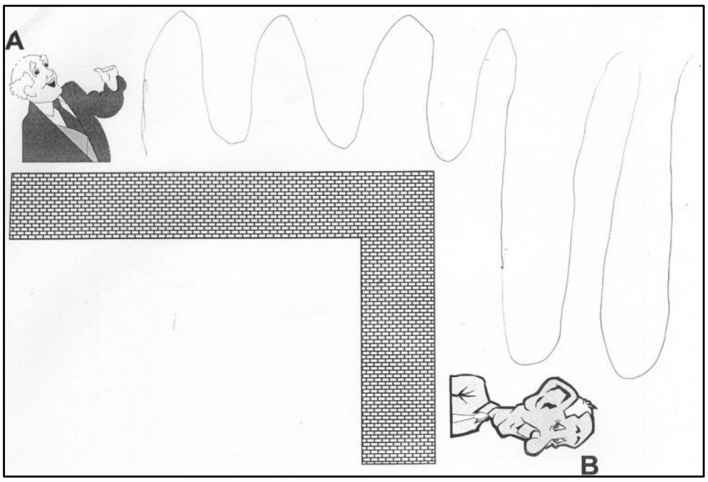


Fig 5.27: Student drawing in response to OES 3 showing waves spreading to fill the space

In OES 3, 11 **rural** students showed sound waves mysteriously bending at 90 degrees, or curving in order to reach the hearer, with no apparent agent for the turning. This is linked to d6 that sound waves cannot be refracted or bent like light, but rural students also didn't show any signs of illustrating refraction, but rather just turned the waves at will as shown in Fig. 5.28 below.

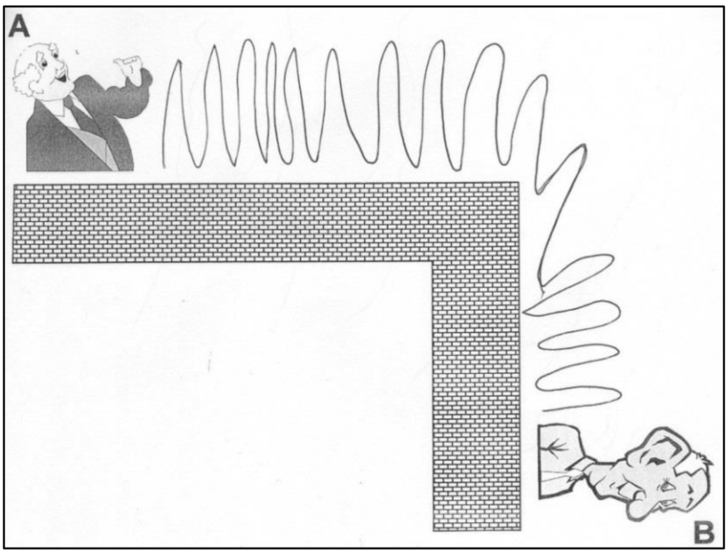


Fig 5.28: Rural student drawing in response to OES 3 wave turning towards hearer

While these questions and the students' answers shed fascinating light on some of their prior mental models and visualizations (C), the limitations on the ER (one 45 minutes show) do not allow for all these concepts to be addressed. The visualization of sound travelling as a sine-wave is deeply rooted and transfers across different contexts, causing further explanatory problems as in this case.

In summary the urban students exhibited some, but certainly not all, of the student difficulties I presented in table 2.2 in section 2.6. As they were Grade 9's, they probably still have their naïve conceptions (or prior knowledge – C) formed from their experience of the world and their perception of how sound waves travel. Their better English language ability helped them compared with the other groups, but they clearly still struggle to express themselves fluently.

In summary the township students exhibited a few but certainly not all of the student difficulties I tabulated. Being also Grade 9's, they probably still have their naïve conceptions (C) formed from their experience of the world and their perception of how sound waves travel. As a group they seldom offered deep explanations and usually reflected their own life-experience rather than the underlying cause. For example in OES 8, only 2 students mentioned vocal cords, while 12 attributed men's deeper voices to puberty or adolescence, as follows:

> (T)" Because it is a natural period of time that a men has change voice because he has been grown up."

Also in OES 9 about hearing sounds in space, students clearly struggled to imagine something they had never experienced:

> (T) "It is because you cannot hear something more what is in earth and also because I have never experienced such thing."

Their moderate English language ability led to a struggle to express themselves fluently. This led to more and more explanations becoming just a repeat of the question, or what was in the voice-bubble: – especially towards the end of the survey

In summary the rural students exhibited *some* of the student difficulties I tabulated, but their very poor English made it difficult to get to the heart of these.

A summary of which student groups showed evidence of each difficulty is shown in table 5.1 below, where "Y" indicates "YES" and "S" indicates "SOMEWHAT". "LEV" refers to the level given to the difficulty in terms of the four-level framework of Grayson, Anderson and Crossley (2001) outlined in section 2.6.

Table 5.1: Summary of student difficulties experienced by each of three student groups

NO	DIFFICULTY	LEV	U	T	R
1	Sound travels as a transverse sine wave	4	Y	Y	S
2	Sound exists as a microscopic entity either carried by or transferred between molecules	4			
3	Sound is a macroscopic entity with impetus, in the form of flowing air or a traveling pattern	4			
4	The speed of a sound-wave is affected by: force, "sound-resistance", separation of molecules	4			
5	Sound can be trapped in a container if the air is trapped-needs small holes to escape	4	Y	Y	
6	Sound can travel through a vacuum and therefore through space	4	Y	Y	
7	Sound waves cannot be refracted or bent like light	3			
8	Confusion between volume and pitch (and amplitude and frequency) and their units	3	Y	Y	Y
9	Sound can only travel if air is present (therefore can't travel through liquids or solids)	3	Y	Y	Y
10	A human has many different vocal cords to produce different sounds	2			
11	Ultrasounds are very loud sounds	2			
12	Sound travels from the hearer to the source (Reverse sound)	2			
13	Music has low volume (small amplitude) and noise has high volume (large amplitude)	2		Y	Y
14	Longer objects vibrate faster, or produce higher notes	2			Y
15	The sound box on a musical instrument is to make the sound clearer	2			Y
16	Vibrations and waves are the same thing	2	Y	Y	
17	Sound travels in electromagnetic Waves	1	Y	S	S
18	Sound is unaffected by solid obstacles, passing right through them	1	Y	Y	S
19	Sound waves turn towards a hearer	1	Y	Y	Y

Key: LEV = Level of difficulty; U = Urban, T = Township, R = Rural Students; S = Somewhat, Y = Yes.

In general it is apparent from table 5.1 that the school groups shared most of the same difficulties, although the rural school was somewhat different in that they showed two difficulties unique to their group (14 and 15) and one shared only with township students (13). Furthermore the three difficulties shown strongly by the other groups (1, 17 and 18) were only weakly evident for the rural group. I did not experience any students displaying d2, d3 or d4, but these are deeper level explanations which would be more likely to come out in interviews than in the simple tests I used. Similarly, I didn't experience d7 either, but again I am not sure if it would have come out in my simple tests. d8 and d9 came out consistently across all groups and one should consider reclassifying these at level 4 (as they also come from the literature). From all the data, I found no evidence of d10, 11 or 12, and indeed have never found any evidence of these while doing this show, so would not suggest raising them to a level 3, and indeed I would sideline them for our local context until I find definite evidence of their occurrence. None of d13 – d16 was found in *all* the groups and I would not want to elevate these to level 3 without finding some more evidence of their regular occurrence, so rather leave them at level 2. The new student difficulties d17 – d19 were fairly consistent across all the schools (although slightly less so for the rural group) and in future studies I would certainly raise them to a level 2. As suggested by Grayson, Anderson and Crossley (2001) in section 2.6, I present a reworking of table 5.1 in the light of my student results (as well as the number of references in the literature and from my experience), in a form more relevant for our local context as table 5.2. The 3rd column shows the new level of each item, compared with the original level on the right hand side.

Table 5.2: Reworked student difficulty table in the light of student data.

NO	DIFFICULTY	NEW LEVEL	U	T	R	NO. OF REFS.	ORIG-INAL LEVEL
1	Sound travels as a transverse sine wave	4	Y	Y	S	9	4
5	Sound can be trapped in a container if the air is trapped-needs small holes to escape	4	Y	Y		4	4
6	Sound can travel through a vacuum and therefore through space	4	Y	Y		4	4
8	Confusion between volume and pitch (and amplitude and frequency) and their units	4	Y	Y	Y	2	3
9	Sound can only travel if air is present (therefore can't travel through liquids or solids)	4	Y	Y	Y	2	3
13	Music has low volume (small amplitude) and noise has high volume (large amplitude)	2		Y	Y	My experience	2
14	Longer objects vibrate faster, or produce higher notes	2			Y	My experience	2
15	The sound box on a musical instrument is to make the sound clearer	2			Y	My experience	2
16	Vibrations and waves are the same thing	2	Y	Y		My experience	2
17	Sound travels in electromagnetic Waves	2	Y	S	S		1
18	Sound is unaffected by solid obstacles, passing right through them	2	Y	Y	S		1
19	Sound waves turn towards a hearer	2	Y	Y	Y		1

Key: U = Urban, T = Township, R = Rural Students; S = Somewhat, Y = Yes.

In conclusion this section yielded some fascinating insights into student difficulties in sound. My students' difficulties were not unlike those in the literature, although it was interesting that a few in table 5.1 were not found in this group, so have been left out of table 5.2. It was also interesting to discover 3 new difficulties not found in the literature and to add these as d 17, 18 and 19. Drawings and written answers (where intelligible) especially provide data towards understanding the R-M factor of the show (the ER) enabling us to consider ways to improve the show. It must be stressed though, that what I was analyzing represents the *expressed* models of the students and more work would need to be done (perhaps through interviews) to gain more access to their personal mental models

5.3 Prior knowledge before, and learning during the show.

This section focuses on whether the show led to any increased knowledge and understanding of sound, and focuses on the multiple choice (MC) pre- and post-test questions (MCQ 1 – 10). Pre-test results reflect on the C-factor of students' prior-knowledge before exposure to the ER. Post-test results reflect on the R-C factor indicating to what extent new knowledge was learnt (accommodated, assimilated and integrated) from the ER.

Urban students displayed weak prior knowledge of concepts in sound and waves as reflected in their low overall pre-test class average of 30 %, but as they were Grade 9's and had not yet covered this material in class, this could be expected. There was good prior knowledge of some of the basic facts like the units of volume and frequency (questions 4 and 5) with more than half of the class knowing these, but not in any of the other questions in the multiple choice. In most of the questions urban students simply did not know the answers, as reflected in the group average being very close to chance (25%) as in questions 6 – 10, or even below, as in questions 1 – 3. The graph below shows the percentage of urban students choosing each option (A – D) with the correct option (or key) coloured in red (Fig. 5.29). The graph clearly shows common wrong answers for questions 1 – 3 and correct answers for questions 4 and 5. For questions 6 to 10 there is no clear pattern and students were probably just guessing.

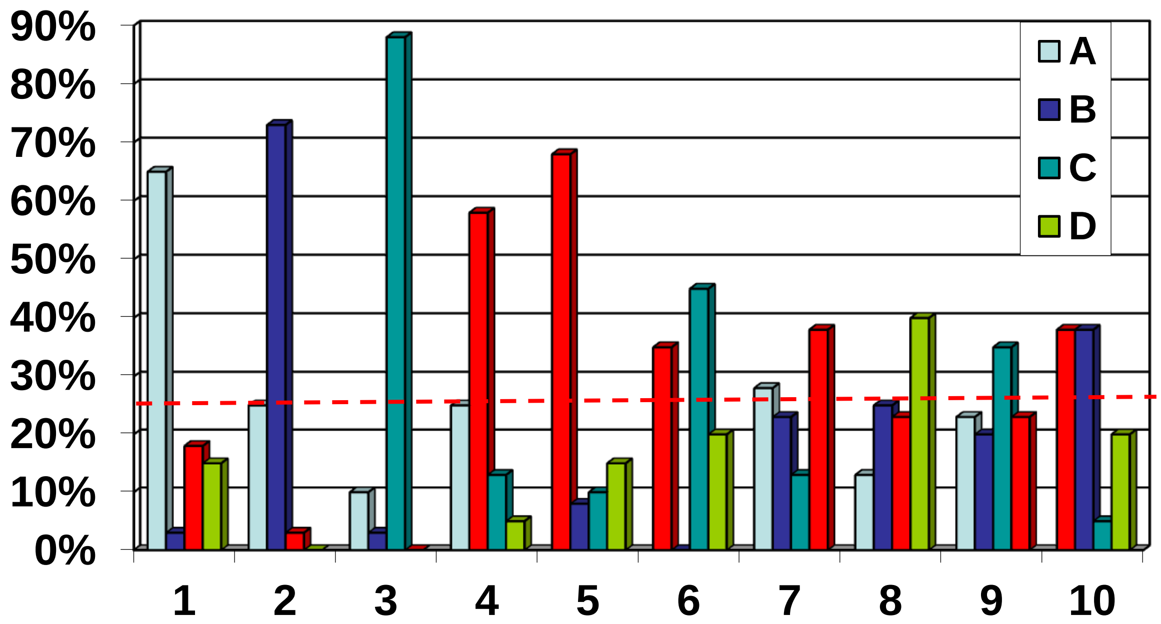


Fig 5.29: Distracter analysis graph for urban pre-test. Percentage of students choosing options A – D for each of questions 1 – 10.

Most of the urban students' naïve conceptions revealed in the MCQ Pre-test were remediated during the show and not evident in the post-test. There was also clear evidence of learning in this group from the MCQ pre-test to the post-test, with the group average almost doubling from 30 to 59 % The data showed an extremely low T-test p-value of 1.08E-11 proving high statistical significance (p much < 0.05). In addition the pre-test showed only four questions where the correct answer was the one most chosen, but in the post-test this was true in eight questions: all except the comparison question,3, and question 1. The student conception in MCQ 1 was that vibrations and waves are the same thing (d 16) and the incorrect answer (a) was not corrected by the show. In contrast, the difficulty (d17,) that sound travels in electromagnetic waves, represented by MCQ 2 (b) was almost completely dealt with by the show (as were weakly held student prior conceptions in MCQ 6, 8, 9 and 10). This is graphically displayed in Figure 5.30 below. The contrast with the graph in Fig. 5.29 is striking.

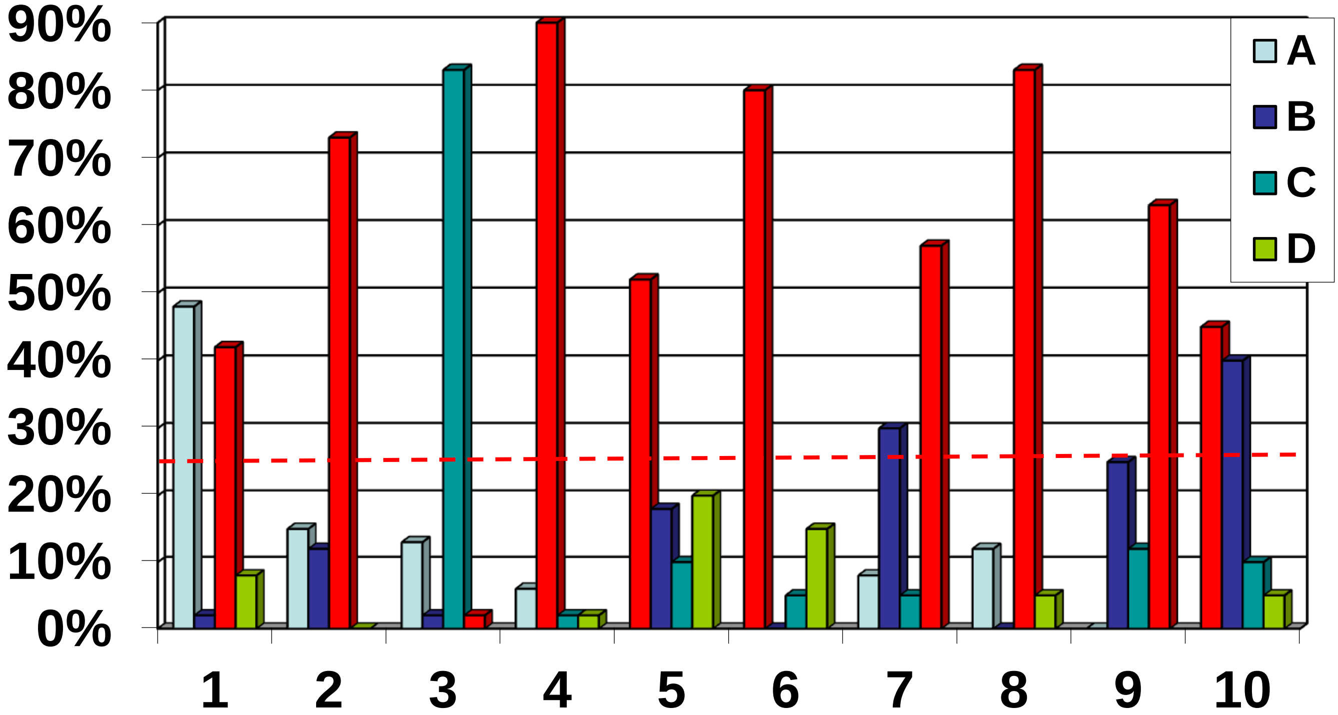


Fig 5.30: Distracter analysis graph for urban post-test. Percentage of students choosing options A – D for each of questions 1 – 10. (Key is coloured red)

In almost every case the correct answer was a clear first choice with over half of the class choosing it. There were three questions (4, 6 and 8) where over 80 % of students chose the correct answer in the post-test (and none like this in the pre-test). Especially in the questions directly related to the 5 key concepts in the show (MCQ 1, 2, 6, 7, 8, 9) urban students fared extremely well and showed dramatic increases from the pre-test. While the longer Open-ended Survey questions were designed more to probe difficulties than learning during the show, some of the sound answers written shown in previous sections and the correct drawings show that learning had taken place.

Township students similarly displayed extremely weak prior knowledge of concepts in sound and waves as reflected in the lowest pre-test score of 18 %, which is very poor even though they were Grade 9's and had not yet covered this material in class. In most of the questions, township students simply did not know the answers, as reflected in the group average for the correct answer being very close to chance (25%) as in questions 1, 2, 4, 5, 8, 9, 10, (in Fig. 5.31) below chance as in questions 6

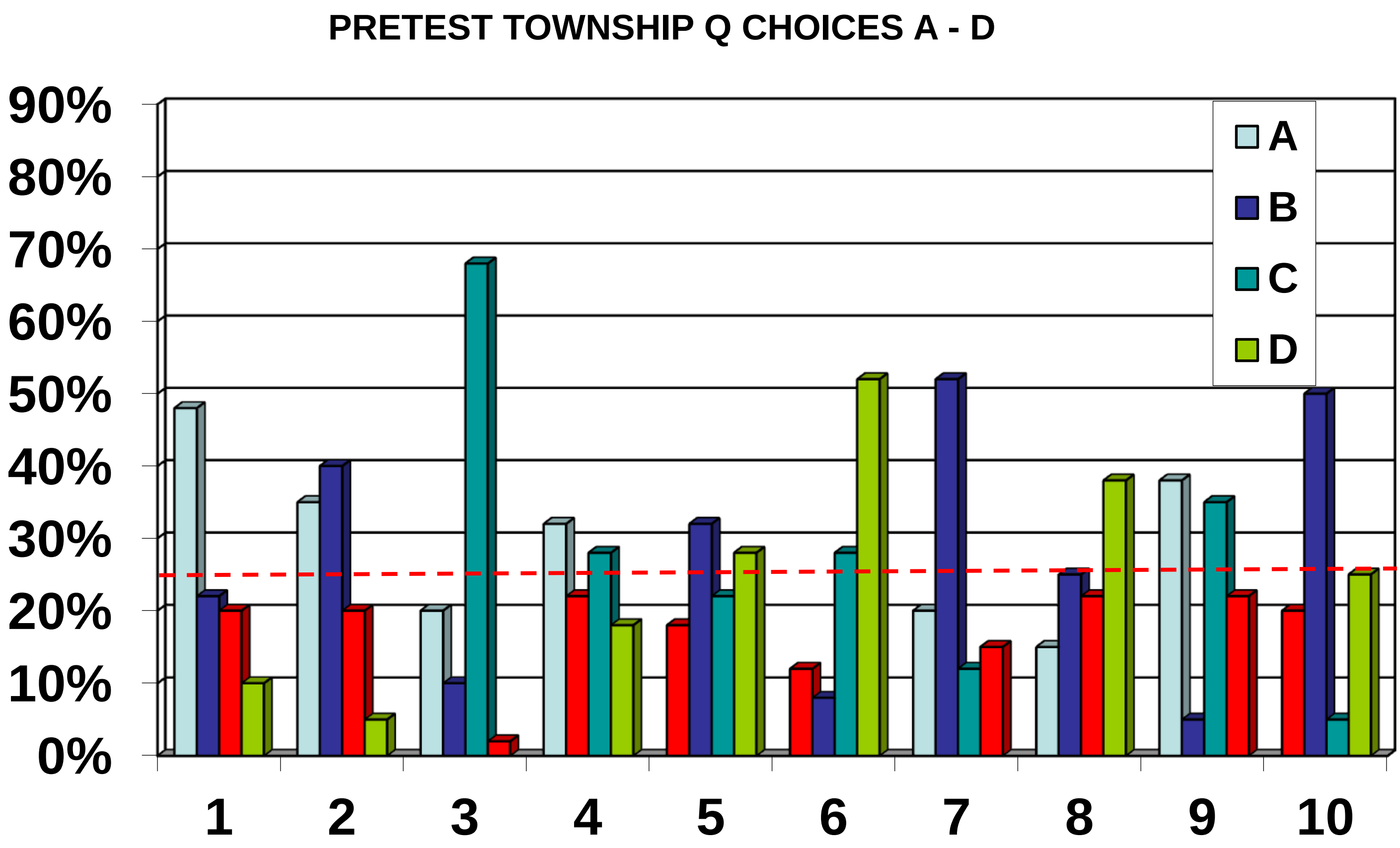


Fig 5.31: Distracter analysis graph for township pre-test. Percentage of students choosing options A – D for each of questions 1 – 10. (Key is coloured red)

and 7; or way below, as in questions 3. In *no* question was the most common choice the correct answer, and township students displayed incorrect prior conceptions in every question.

Many of the township student conceptions revealed in the MCQ pre-test and mentioned above were sorted out after the show and not evident in the post-test. There was clear evidence of learning in this group from the MCQ pre-test and post-test, with the group average more than doubling from 18 to 38 %, and a very low T-test p-value of 9.18 E -09 proving high statistical significance (p much < 0.05). In addition the pre-test showed no questions where the correct answer was the one most chosen, but the post-test showed six questions where this was true . These 6 were in fact the questions directly related to the 5 key concepts in the show (MCQ 1, 2, 6, 7, 8, 9) where township students fared extremely well and showed dramatic increases from the pre-test. Post-test results are shown in Fig. 5.32 below. Apart from the comparison MCQ 3, Township students showed confusion between volume and pitch (d 8), in terms of their units, in MCQ 4 and 5 which were not improved after the show. The confusion between amplitude and frequency (also in d 8) was evident in the pre-test (MCQ 6 and 7) but corrected in the post-test. Their conception that sounds which humans cannot hear (but dogs can) have a very low frequency; answer (b) for MCQ 10, was also not fully corrected by the show (although there was some improvement visible).

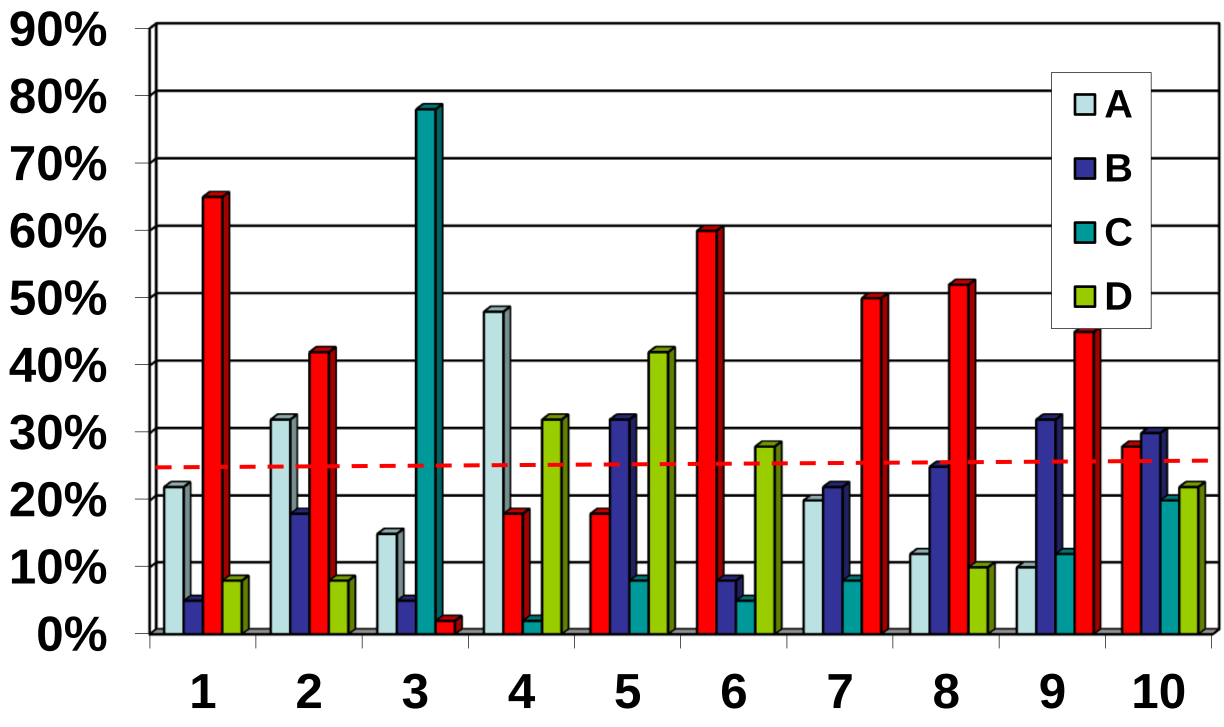


Fig 5.32: Distracter analysis graph for township post-test. Percentage of students choosing options A – D for each of questions 1 – 10. (Key is coloured red)

Like the township students, **rural** students displayed extremely weak prior knowledge of concepts in sound and waves as reflected in their very low pre-test score of 19 %. Although they were Grade 9's and had not yet covered this material in class, this score was less than one would get just by guessing (25%) indicating some major misunderstandings. The group started well with 38% of the class answering the first question correctly, and 41 % answering question 6 correctly. Question 1 increased to 60 % in the post-test, but this was the only question showing such an increase. For the rest of the questions rural students simply did not know the answers, as reflected in the group average for the correct answer being less than chance (25%) in every other case. See Fig. 5.33 below. Apart from the comparison question 3, there was little evidence of strong prior conceptions and the graph indicates mostly guesswork. MCQ 4 and 5 were both incorrectly answered with (c) metres per second as the unit of volume and frequency respectively. This was not really cleared up after the show as MCQ 4 showed no change and MCQ 5 showed a change to (b), decibels, still showing evidence of confusion between volume and pitch and their units (d 8). This was also evident in answers to MCQ 6 and 7 (although MCQ 7 did see some improvement). MCQ 9 and 10 showed interesting trends which will be discussed after this section.

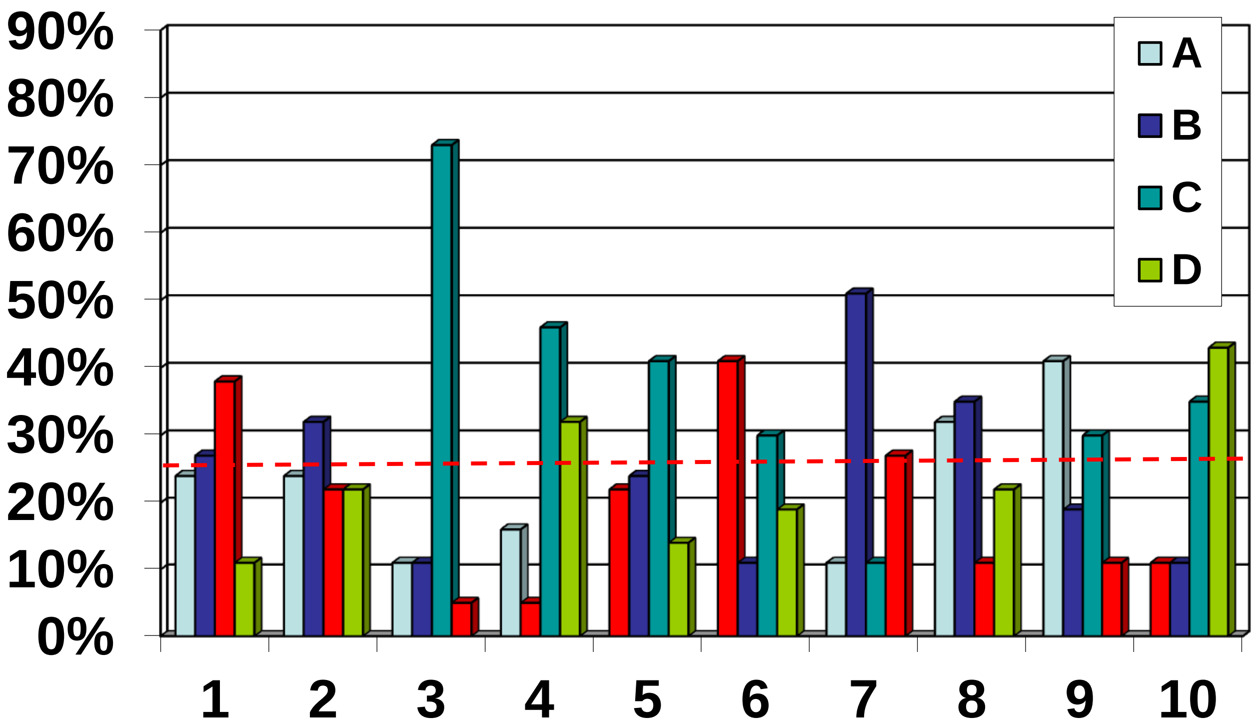


Fig 5.33: Distracter analysis graph for rural pre-test. Percentage of students choosing options A – D for each of questions 1 – 10. (Key is coloured red)

There was little evidence of learning in this group from the MCQ pre-test and post-test, with the group average changing only from 19 to 27 %, although a low T-test p-value of 0.00159 (< 0.05) still shows statistical significance. When considering individual questions, it is even clearer that little learning took place during the show. What is striking about the pre- and post-test graphs is how similar they appear. In the pre-test, there were two questions where the correct answer was the one most often chosen (questions 1 and 6), and in the post-test still only two questions where this was true (question 1, again, and 7). Especially in the questions directly related to the 5 key concepts in the show (MCQ 1, 2, 6, 7, 8, 9) rural students fared poorly and showed little improvement from the pre-test. It was also extremely difficult to gauge from the written answers if learning had taken place, as the standard of their English was so poor. See Fig. 5.34 below:

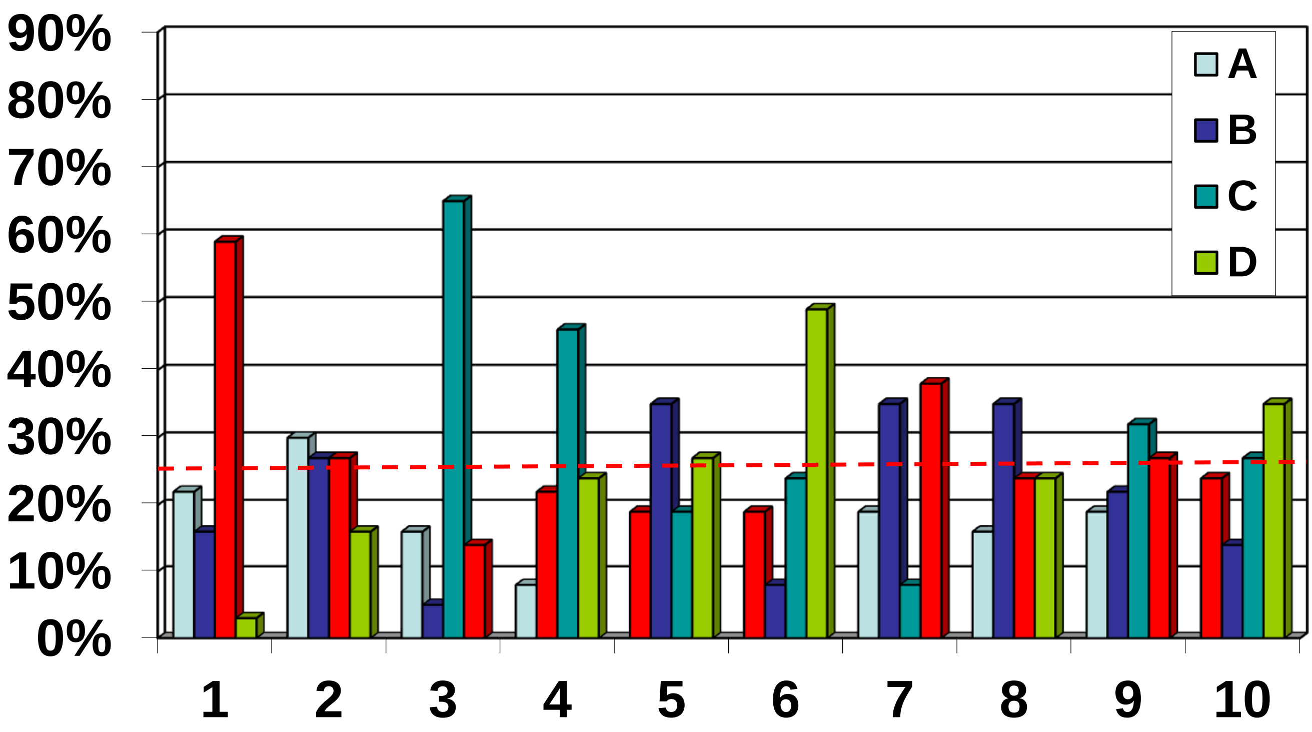


Fig 5.34: Distracter analysis graph for rural post-test. Percentage of students choosing options A – D for each of questions 1 – 10. (Key is coloured red)

In conclusion, let us consider some results which were evident across all groups. It has already been stated that all groups shared the common difficulty (d 9) that sound travels fastest through air (rather than solids) and that none of them changed this after the show. This was chosen as a comparison question as it represents a well-documented difficulty (Caleon and Subramaniam (2010), Linder (1992)) and was not taught in the show in any way. Consequently the incorrect option (c) was chosen by the majority of students across all groups both in the pre- and the post-tests. Let us also compare the different groups' answers for the last two questions (MCQ 9 and 10). MCQ 9 asked students to choose the correct answer (underlined below) for the following:

9) A *longer* guitar string will:

a) vibrate faster, and produce a higher note than a shorter string
b) vibrate faster, and produce a lower note than a shorter string
c) vibrate slower, and produce a higher note than a shorter string
<u>d) vibrate slower, and produce a lower note than a shorter string</u>

The graphs below show the distractor analysis for the pre-test (figure 5.35) and post-test (figure 5.36) for MCQ 9 only, for the three groups and all groups together. In the pre-test the pattern is not very clear when looking at individual groups. When considering all groups together, the two most popular options in the pre-test are (a) and (c), both of which (incorrectly) state that a higher note will be produced, listed as difficulty d 14 on table 2.2. One can see that this more common in the township and rural groups, and somewhat in the urban group. In the post-test the most popular answer is the correct one, (d), followed by (b), both of which (correctly) state that the note produced by a longer string will be lower. As this concept (KC 5: "long is low, short is high") was repeated many times and demonstrated with many different instruments during the show, it is pleasing to see this change for the groups taken together. Looking at individual groups, urban and township students clearly made the correction after the show, but with the rural students there is still a lot of confusion and strong evidence of this difficulty (d14), although there was at least a measurable increase in students selecting the correct answer.

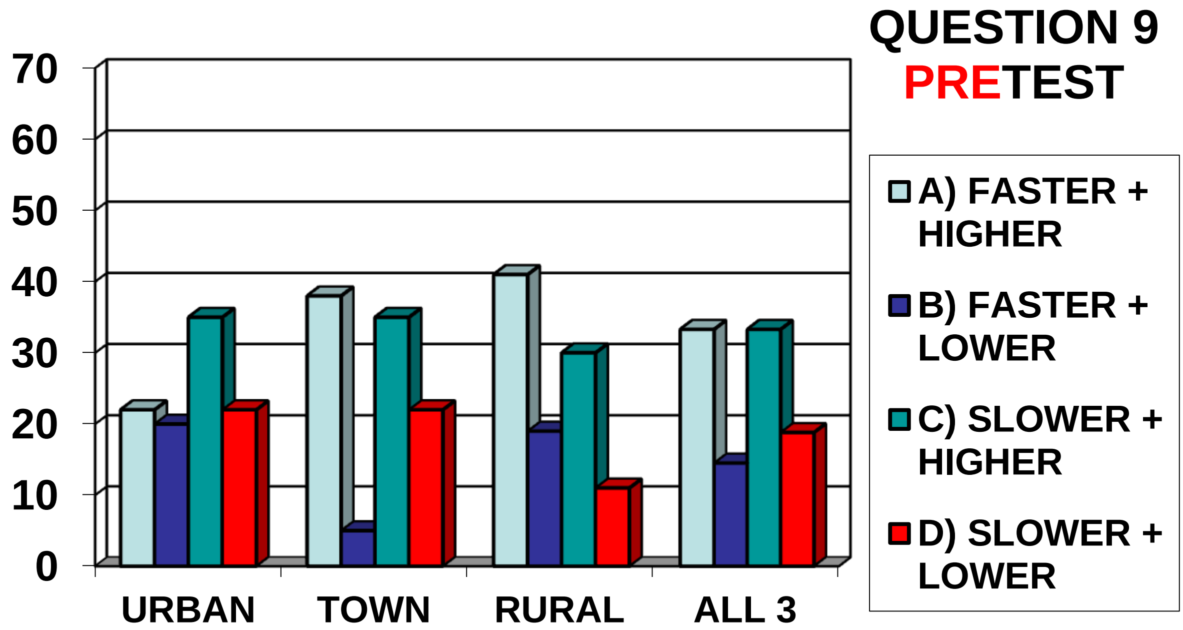


Figure 5.35 Percentage of students choosing each option for MCQ 9 in the pre-test

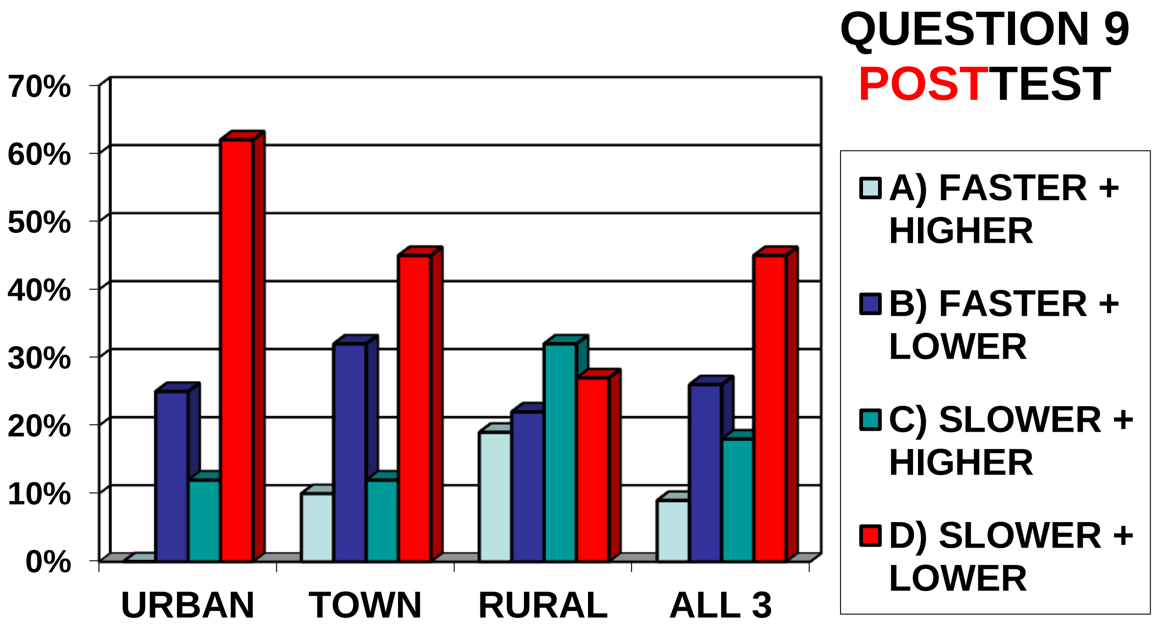


Figure 5.36 Percentage of students choosing each option for MCQ 9 in the post-test

MCQ 10 asked students to choose the correct answer (underlined below) for the following:

10) Dogs can hear some sounds which humans cannot hear, because these sounds:

a) have a very high frequency

b) have a very low frequency

c) are too loud

d) are too soft

The graphs below show the distractor analysis for the pre-test (figure 5.37) and post-test (figure 5.38) for MCQ 10 only, for the three groups and all groups together. When considering all groups together, the most popular options in the pre-test are not too clear and all four options were selected by quite a few students. It is interesting to look at the differences in the pre-test selections for each of the groups though. Urban students most commonly chose options (a) and (b), both of which indicated that the reason had to do with frequency, although they were not too sure if it was high or low frequency. These two frequency options were still their most popular choices in the post-test, with the correct answer only slightly more popular. Rural students, in contrast, chose options (c) and (d), both of which indicated that the reason had to do with volume (loud or soft), although they were not too sure if it was too loud or too soft. These two volume options were still their most popular choices in the post-test, with selections of the correct answer increasing, but still not as popular.

Township students most commonly chose options (b) and (d), both of which indicated that the reason had to do something being too low (frequency or volume), with low frequency (b) being their more popular choice. The incorrect low frequency option was still their most popular choice in the post-test, with the correct answer becoming more popular. For all three groups the correct answer (a) did finally emerge as the most popular, but with the three distractors each being selected by at least 20% of the students. The differences between groups are interesting: it appears that urban students are more familiar with the concept of frequency, and they had no problems with MCQ 5 or 7 in either the pre- or the post-test. In contrast rural students did not know the unit of frequency (MCQ 5) in either the pre- or the post-test and struggled with MCQ 7 in the pre-test. Their response shows up the confusion between volume and pitch (and amplitude and frequency) which is listed as difficulty d 8 on table 2.2.

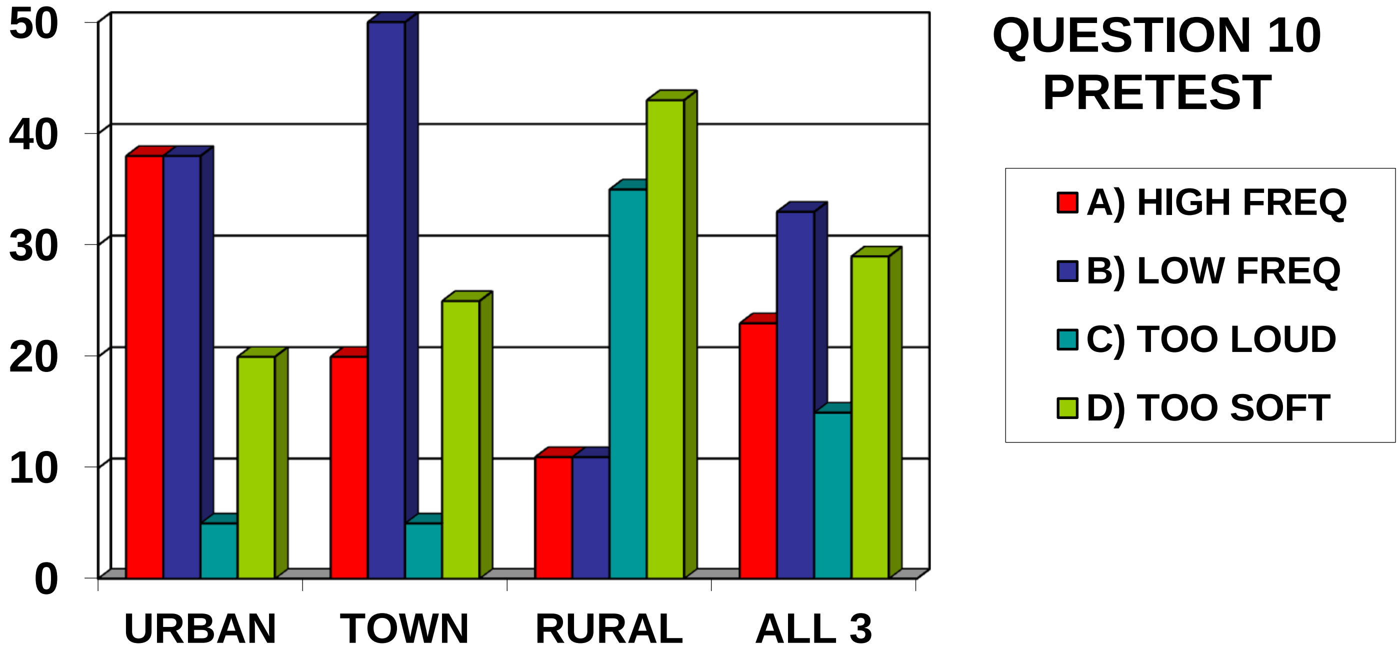


Figure 5.37: Percentage of students choosing each option for MCQ 10 in the pre-test

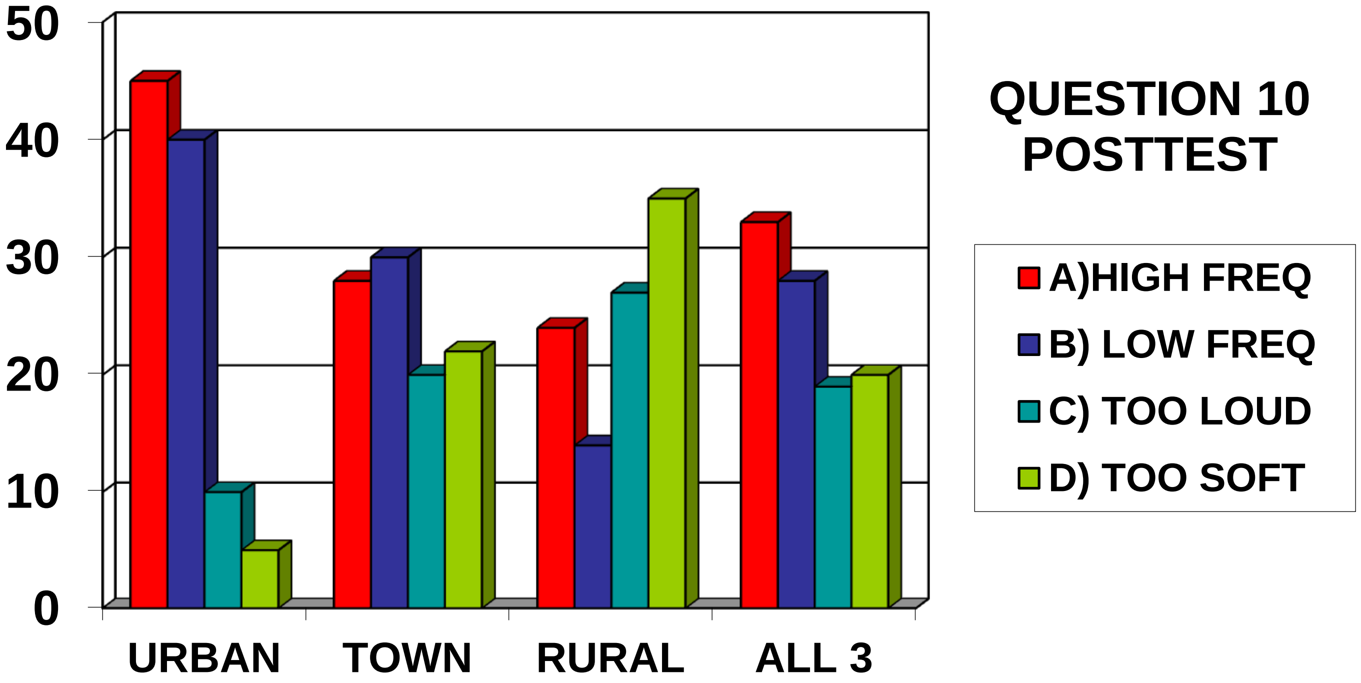


Figure 5.38: Percentage of students choosing each option for MCQ 10 in the post-test

5.4 Summary and Conclusions: Group Performance.

In general the **urban** students performed better than the other two student groups in every department. Apart from the pre- and post-test scores mentioned above, in the multiple-choice, "Choose an opinion" section, they averaged 19.4 students correct out of 40, or just less than half. In the open-ended survey, their written explanations and their drawings showed some depth and looked for underlying causes for the effects being questioned. They were the *only* group to commonly mention (or draw):

- Particles or atoms when explaining speed of sound in different phases (OES 5 – see Fig. 5.39 below)
- The effect on the ear-drum when explaining why loud sounds hurt our ears (OES 7)
- The changes in the vocal cords when explaining why men have deeper voices than ladies (OES 8)
- The lack of air in space when explaining why sound doesn't travel through space (OES 9)

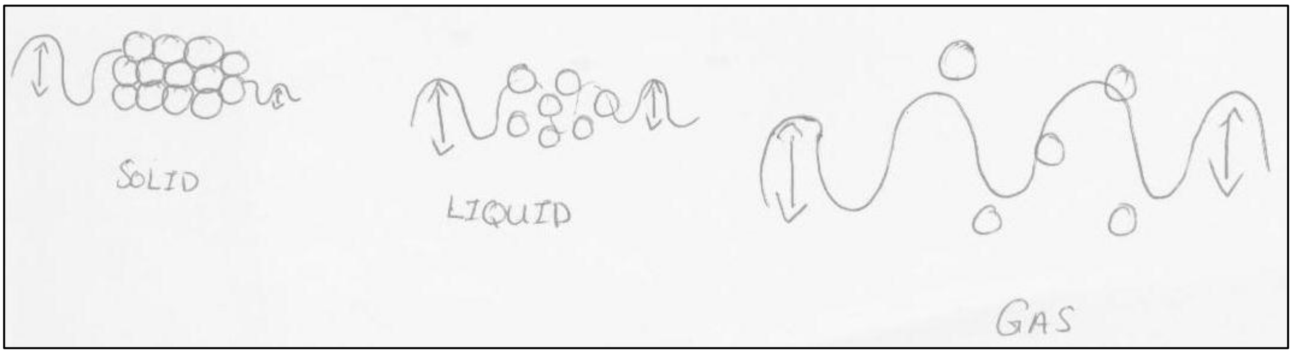


Fig 5.39: Urban student drawing in response to OES 5 showing consideration of particles

In general the answers from the urban group showed the ability to abstract and to think around an issue rather than merely describing what one has experienced. They frequently described phenomena from a wave-based perspective rather than an object based one. The very high group average for the English Proficiency Test (81.8 %) confirms that they had a distinct advantage in this area over the other two groups which assisted them in writing the tests but even more so in hearing, understanding and learning from the show. This shows the advantage of attending an urban school, as most of the students do not speak English as a first language (but rather isiZulu) but have gained good English skills from their time at the school. They were the only group who scored well on the C-R-M factor as they were able to interpret, visualize and learn from the ER fairly successfully. Although they started with weak prior knowledge (C), their MCQ post-test showed good ability to assimilate the new knowledge into their existing schemes (R-C), improving on almost all the questions, except the comparison MCQ 3 and MCQ 1. In MCQ 1 they expressed the incorrect prior conception that vibrations and waves are the same thing (d 16) and their incorrect answer (a) was not corrected by the show. Although they were confident to attempt drawings and showed some visualization skills (R-M) there was still much confusion between representation and reality

Urban students also showed good recall of what they had seen in the show, even after six months, with most of the 10 students I interviewed recalling at least five of the demonstrations they saw in the show at a superficial level. Among the instruments they saw, they recalled especially the drums, marimba and harmonica (4 urban students remembered the harmonica, whereas none from any of the other schools did). To a lesser extent, they recalled the guitar, kazoo, wooden frog piston flute and French Horn (with no other school remembering this instrument). None remembered the piano accordion, which was quite well remembered by the other schools. Nine out of ten students recalled the waves they had seen on the screen and were able to remember some of the activities done with these waves – far more than in any other group. Clearly the technological aspects of this show resonated with this group. Surprisingly to me, no students mentioned the unusual instruments which I had thought would stick in their memory: the nose flute, hose trumpet and whirly pipe.

Township students started at an even lower level of prior knowledge (C) but showed very pleasing improvement in the post-test (R-C), like the urban groups seeming able to assimilate new knowledge without too much stress. Questions which did not improve and still showed evidence of difficulties included MCQ 4 and 5 (d 8) and MCQ 10, as reported earlier. In the multiple-choice, "Choose an opinion" section, they averaged 18.4 students correct out of 40, or just less than half, and very comparable with the urban school. Their written answers were moderate, but many students left the space blank or just rewrote the question (or the information from the "voice-bubbles"). In general their written explanations showed weak English skills and lacked depth. They mostly gave explanations in terms of life-experiences and rarely mentioned the underlying causes for the effects being questioned. Their medium group average score (66%) for the English Proficiency Test (4.3) confirms that they have *some* problems in this area, both in writing the tests but also in hearing, understanding and learning from the show. The fact that their multiple choice tests improved greatly, but their later written answers were weak, seems to indicate that they can hear and read English satisfactorily, but struggle to express themselves when writing. Thus there is potential for learning during a show like this, and for detecting that through a multiple choice test, but a written test may not prove to be as useful as an indicator of learning. In general, reasoning with the ER (R-M) was not strong as indicated by these weak written answers and drawings which were more illustrations than explanations.

Township students also showed good recall of what they had seen in the show after six months with most of the 10 students I interviewed remembering at least five of the demonstrations they saw in the show. Among the instruments they saw, they recalled especially the guitar, drums and kazoo. To a lesser extent, they recalled the piano accordion, vuvuzela and "drumskin whirler". (They were the only school to remember the accordion, perhaps because Vusi Ximba, a township artist made it popular?). Six out of ten students recalled the waves they had seen on the screen – so the technological aspects of this show resonated somewhat with this group. Three of these students mentioned the example given of the different lengths of violin, viola, cello and bass, which few others remembered.

Rural students started from a very low level of prior knowledge (C) as demonstrated by their very poor scores on the pre-test. They therefore struggled to make much headway in improving their knowledge, as supported by the data in the previous section. Their post-test MCQ scores showed hardly any improvement and it seemed that accommodating new knowledge was stressful for them (R-C). Especially in the questions directly related to the 5 key concepts in the show (MCQ 1, 2, 6, 7, 8, 9) rural students fared poorly and showed little improvement from the pre-test. In the multiple-choice, "Choose an opinion" section, they averaged 14 students correct out of 37, the lowest if all the groups. In general the few open-ended survey written answers from the rural group which were useful lacked depth and the ability to abstract and to think around an issue rather than merely describing what one has experienced. One must exercise caution here though as this study is not sufficient to explore the full reasons for their weak answers. The influence of English language ability is naturally huge, but Lee (1998) has pointed out that students from cultures which do not encourage longer or more detailed responses can be perceived as lacking full and complete knowledge. This is the value of triangulating the results of multiple probes for these students. When rural students did offer an intelligible explanation, it was usually based around a life-experience or an alternative worldview, as similarly observed by Lee (1999) in different contexts. This can be extremely limiting if their understanding of science is constrained by what they have personally experienced. They tended to describe phenomena from an object -based perspective rather than a wave-based one, and the word "wave" was used far less than with the other groups.

The rural group was extremely unwilling to attempt drawings and those they did present were again more illustrations than explanations. It appears they struggled to reason with the ER itself (R-M) and to make sense of (or to redraw) many of the visualizations, pictures and models.

The lowest group average (54%) for the English Proficiency Test (appendix k) confirms that the rural group has severe challenges in this area. This not only affects them adversely in writing the tests, but more so in hearing, understanding and learning from the show. This made it clear that the major factor affecting these students is proficiency in English, and that they still need instruction in isiZulu. These results also bear testimony to the fact that this group is severely disadvantaged in a written exam – even a multiple choice one. The spread of choices in the pre-test graph seem to indicate a random guessing more than a careful selection of the correct answer. That this spread is still evident in the post-test graph indicates that they have learnt little, and still struggle with this type of test. Interestingly, this is exactly the type of test they are forced to write in the South African matric exams, hence their school's 2015 record of a 30 % pass-rate).

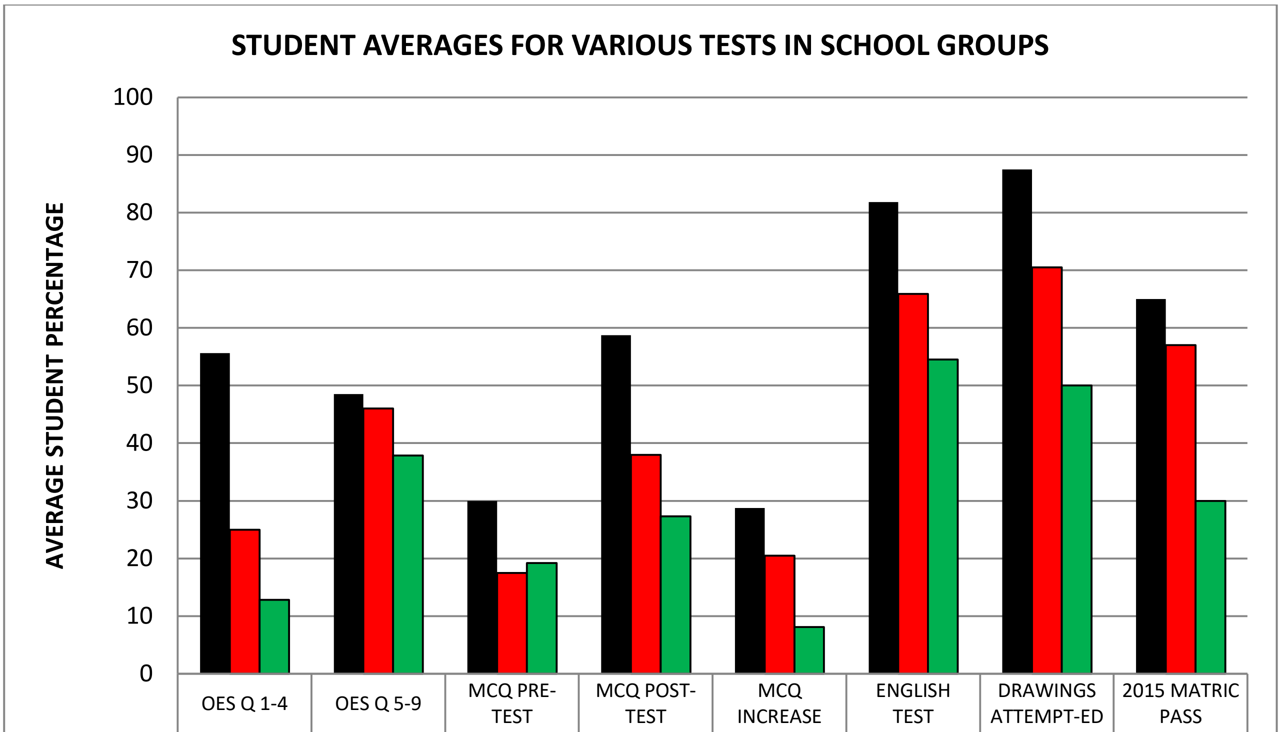
Rural students also showed good recall of what they had seen in the show after six months, with most of the 10 students I interviewed remembering at least five of the demonstrations they saw in the show, although time did not allow me to probe just how much they remembered of the demonstrations. Among the instruments they saw, they recalled especially the guitar, drums, kazoo and vuvuzela. To a lesser extent, they recalled the marimba, wooden frog, piston flute and ingungu (a traditional Zulu instrument). Only one remembered the piano accordion, which was quite well remembered by the township school. Only three out of ten students recalled the waves they had seen on the screen – far less than in any other group, so clearly the technological aspects of this show did not resonate much with this group. Definitely the instruments which were most familiar to them from their homes and culture were best remembered, and this should be remembered along with Aikenhead's (1996) assertion that science learning occurs best when linked to the everyday worlds of the students. What stood out with rural students is that they all had firm memories of the event itself, which they offered without prompting. They mentioned using the clickers writing a test, and especially getting a free cool drink and pen. Obviously the event evoked good memories for them even after a long time. This makes it important for me to ensure that visitors have a good experience at Unizulu Science Centre.

Some general comparisons from the open-ended survey questions: The three groups provided very different data in terms of their drawings, even just in their willingness to attempt them, as detailed in 4.2.1, with half of the rural students leaving out the drawing for each question on average, as opposed to only 13 % of urban students (and 30 % for township students). In addition, only urban students in general provided drawings which attempted to convey the underlying reason for the phenomenon. Those from the township and rural groups were mostly more illustrations than explanations. It is possible, though, that student drawings represent what was convenient or what they felt confident to draw, rather than strongly held misconceptions. In terms of their reported source of knowledge: rural and urban students each reported a 40: 60 % split for prior knowledge versus knowledge gained in the show. With township it was higher at about 30: 70 %. It is pleasing to see that students considered themselves to have learnt new things in the show and the fact that results were much higher for the (multiple choice) post-tests than for the pre-tests bears this out.

For the "choose an answer: multiple choice style part of the open-ended survey, results echoed the findings of the multiple choice post-test, that urban students performed better than township, who in turn were better than rural students, although the differences were not as marked. Finally, the written answers were very disappointing generally and the township and rural answers were so poorly written as to be of little use. Students across the board were clearly not practised or skilled in explaining themselves in writing. An interesting fact which emerged very clearly was that urban students attempted to get to the root cause when offering explanations, and often looked "below the surface" in discussing vocal cords or particles. Township and rural students tended to explain things more in terms of their own life experiences. Table 5.3 below summarises the main results for each group from 5.1 to 5.3 and figure 5.40 depicts this graphically.

Table 5.3: Summary of main results for each school group in original study

SCHOOLS	URBAN	TOWN	RURAL
BIOGRAPHICAL DATA			
Number of students	40	40	37
Black students (isiZulu speaking)	30	40	37
Other students	10	0	0
Female	28	20	19
Male	12	20	18
Average age	15.1	15.3	15.6
RESULTS			
OES % correct written Q 1-4	55.63	25.00	12.84
OES % correct Q 5 - 9	48.50	46.00	37.84
MCQ Pre-test %	30.00	17.50	19.19
MCQ Post-test %	58.75	38.00	27.30
Increase: Pre-post %	28.75	20.50	8.11
English Test %	81.80	65.90	54.50
% Drawings attempted	87.50	70.50	50.00
2015 Matric science pass rate	65 %	57 %	30 %
SOURCE OF KNOWLEDGE			
P – known before show	106	63	86
S - learnt after show	167	166	120
Before/ After	0.6	0.4	0.7
LANGUAGE/ EXPLANATIONS			
Language	Weak but OK	Very weak	Almost unintelligible
Explanations	Tried to identify root cause	Experiential, life-experience explanation	UNABLE TO EXPLAIN. Repeated question or voice bubbles
Scientific nature of explanations	Wave-based	Blend or Object based	(Object-based) when intelligible
MEMORY			
Unique Memory after 6 months	Computer waves on screen	Piano accordion and instruments	The experience: test, pen and cool drink



	OES Q 1-4	OES Q 5-9	MCQ PRE-TEST	MCQ POST-TEST	MCQ INCREASE	ENGLISH TEST	DRAWINGS ATTEMPT-ED	2015 MATRIC PASS
URBAN	56	49	30	59	29	82	88	65
TOWNSHIP	25	46	18	38	21	66	71	57
RURAL	13	38	19	27	8	55	50	30

Figure 5.40: Graph of student percentages for various tests in school groups

The table on the previous page (table 5.3) illustrated in the graph above (Fig. 5.40) shows the urban group being ahead in every single test I administered, in many ways a mirror of their school's pass-rate in the 2015 matric science paper (right hand column). In conclusion it can be seen that the three groups differed vastly in terms of their performance. In every test shown above the urban school outperformed the township school that beat the rural school. This highlights important differences both in their prior knowledge (C) and in their ability to learn from the show (R-C) and must be carefully considered when dealing with the different groups. The data in this section make it clear that one cannot merely assume that each group has successfully interpreted, visualized and learnt from the ER (C-R-M).

Chapter 6: Redesign of the show (Stage 4)

6.1 Introduction and summary of student difficulty data

Chapter 5 reports clear evidence of learning during the show especially for urban and township schools (with each doubling their pre-test scores in the post-test showing promise with the R-C factor) but limited learning for the rural group. In addition clear visualization difficulties were highlighted (R-M) for all groups especially in the drawing of sound waves. Nevertheless it is possible that this learning can be improved, especially where nature of the show itself (M) may be the cause of misunderstanding (as reported in section 5.2, d1). Thus on the basis of all stated above, efforts were made to improve the show.

The four-stage process shown in figure 2.5, (from Trujillo, Anderson and Pelaez, 2016) and mentioned in section 2.7 was presented as a useful framework for the process of developing a science show. At this point in this study I now move to stage 4: redesign: looking to improve both the show (chapter 6) and the testing instruments (chapter 7) in the light of data gathered during stage 3. This will lead to a reiteration of stage 2 (improving the "classroom model") and stage 3 – retesting with consequent new results (chapter 9) after performing a validation exercise (chapter 8). A final redesign of the show by translating it into isiZulu will lead to a third iteration of testing (Chapter 10). Finally, I shall make practical recommendations of how these iterations can ultimately assist science centre practitioners in designing, testing and improving science shows (chapter 11).

To this end, I shall address the two final research questions, repeated here for ease of use:

> 3) Building on the data obtained from addressing RQ 2, how can the show be redesigned to improve understanding and minimize conceptual and visualization difficulties with sound?
> 4) Building on the data obtained from addressing RQ 1, how can the probes be modified so as to be more effective and valid tools for revealing evidence of student learning and any related conceptual and visual difficulties after experiencing *the sound show*?

This chapter will address RQ 3: How can the show be adjusted in the light of student results to provide more understanding across all student groups (subject to the constraints of time and circumstance). Chapter 5 provides comprehensive student results and discussion from the original survey. Table 5.1 summarises the student difficulties encountered and persisting even after the show in terms of the original listing in table 2.2 and presents three novel difficulties not previously encountered. This led to a condensed, reworked table 5.2 in the light of this student data. This table will be used as a basis for adjustments to the show. As discussed in section 5.2, student data reported in table 5.1 showed no evidence of difficulties 2, 3, 4, 7, 10, 11 or 12. In redesigning the show these will thus not be considered. In addition, two of our novel difficulties encountered, 18 and 19, were not really covered in the music show so I would not expect them to be affected even if the show were changed (hence they were also left off the table for modifications to the show). Table 6.1 below reflects table 5.2 in indicating which difficulties were still present after the show and in which groups especially. In addition the questions covering that difficulty are listed as well as a suggested modification, fuller details of which are given after the table.

The difficulties suggested for tackling in the table also represent those most relevant to the current CAPS science curriculum which covers sound at Grade 10 and Grade 4 level, as outlined in section 3.1.3. In addition they reflect the 5 key concepts around which the show is built and which were enumerated in section 3.2. These concepts, and the difficulties most closely connected to them, are listed again below.

1. Sound travels in longitudinal waves, which start with a vibration. (d1, d16, d17)
2. The height of the wave (amplitude) indicates the intensity or volume of the sound. (How loud or soft) (d8)
3. The width of the wave (frequency) indicates the pitch of the sound. (How high or low it is) (d8)
4. Music has a repetitive, recognizable wave shape. Noise does not. (d13)
5. “Long is low, short is high.” A vibrating object which is longer vibrates more slowly (at lower frequency) giving a lower pitch. (d14)

Table 6.1: Student difficulties after the show, associated questions and suggested remedies.

NO	DIFFICULTY (Propositional knowledge)	NEW LEV	U	T	R	Q	Modification
1	Sound travels as a transverse sine wave (Should be) LONGITUDINAL – major issue – dealt with in intro to show	4	Y	Y	S	OES 1,2,3 MC 2	Taught specifically – slinky demo; Changed slides Showed diagrams & animations
5	Sound can be trapped in a container if the air is trapped-needs small holes to escape SOUND TRAVELS THROUGH WALLS	4	Y	Y		OES 2	Shown not to be true with DEMO Cellphone in sealed lunchbox
6	Sound can travel through a vacuum and therefore through space SOUND NEEDS A MEDIUM – stressed in introduction to vibration types	4	Y	Y		OES 9	Show not to be true with VIDEO E.g. bell in evacuated jar
8	Confusion between volume and pitch (and amplitude and frequency) and their units Important difference – language issue – care was taken with words!	4	Y	Y	Y	MC 6, 7	Chose words carefully Taught explicitly. Sang or shouted words
9	Sound can travel only if air is present (therefore can't travel through liquids or solids) TRAVELS FASTEST THRU SOLIDS – major issue – but deliberately not taught to serve as a comparison question	4	Y	Y	Y	OES 5 MC 3	NO change
13	Music has low volume (small amplitude) and noise has high volume (large amplitude) MUSIC HAS SMOOTH REPETITIVE WAVE Important difference – language issue – care was taken with words!	2	Y	Y	Y	OES 4 MC 8	Chose words carefully Taught explicitly when doing Data Studio demo
14	Longer objects vibrate faster, or produce higher notes LONG IS LOW, SHORT IS HIGH	2	Y	Y	Y	MC 9	Stressed and explained "long is low" Used multiple instrument examples

	– major issue – demonstrated and explained						
15	The sound box on a musical instrument is to make the sound clearer MAKES SOUND LOUDER – demonstrated and explained	2			Y	OES 6	Chose words carefully Taught explicitly
16	Vibrations and waves are the same thing NOT THE SAME – made clear in intro – showed vibrations causing waves	2	Y	Y		MC 1	Chose words carefully Taught explicitly Demo: ruler and slinky
17	Sound travels in electromagnetic Waves LONGITUDINAL – EM = LIGHT: Covered in introduction	2	Y	S	S	MC 2	Slide on types of waves

6.2 Elucidation of difficulties and modification of show.

D1) Sound travels as a transverse wave (Should be - Longitudinal): This was specifically taught with slides in the introduction, and demonstrated with a "slinky spring". In addition, multiple misleading slides identified in section 5.2 were corrected. Where representations (graphs) of sound waves were retained, they were clearly labelled: "graph of sound wave". (Thus addressing C-M and R-M factors)

D5) Sound needs small holes to escape a container (Sound can travel through walls): This was corrected by a demonstration that a cellphone in a sealed plastic box can be heard (as long as there is air in the box). (R-M and R-C)

D6) Sound can travel through a vacuum and space (Sound needs a medium like air to travel through): This was demonstrated with a video of a bell ringing in a container from which the air is sucked out. (Possibly as a 2^{nd} stage to the demonstration in D5 above) (R-M and R-C)

D8) Confusion between volume and pitch [and amplitude and frequency]. This was previously taught with the help of the Data Studio "Waveport" software, but was now taught more carefully, using only the words "loud and soft" for volume and "high and low" for pitch. In addition the words were made clearer as follows: (C-M and R-M)

- “loud” – shouted
- “soft” – whispered
- “high” – sung at a high pitch
- “low” – sung at a low pitch

D9) Sound travels only through air, not liquids or solids (Sound travels through all media and fastest through solids). This was previously left out of the show deliberately so as to serve as a comparison question as to whether we could link learning with the show or not. This was left out deliberately again. (Thus no attempt was made to address C-R-M factors here)

D13) Music has low volume and noise has high volume (Music has a smooth, repetitive waveform). This was previously taught with the help of the Data Studio “Waveport” software, but was now taught more carefully. (R-M and R-C)

D14) Longer objects vibrate faster producing higher notes (“Long is low, short is high.”). This was demonstrated using multiple instruments (marimba, kalimba, guitar, orchestral strings, piston flute, and straw oboe). In addition care was taken to make “high and low” clear by singing as described above in D8. (C-M and R-M)

D15) The sound box on an instrument is to make the sound clearer (it is to make the sound louder). This was previously clearly demonstrated with a music box on a guitar and the confusion is most probably caused by language, as discussed in sections 4.2.3 and section 7.3 question 6. The demonstration explanation was now carefully worded as per the rewording of Q6 in the open-ended survey to make a clear distinction between: “louder, more distinct or clear, and higher in pitch.” (C-M and R-C)

D16) Vibrations and waves are the same thing (A vibration causes a wave and is not the same thing. A vibration is a repetitive movement in one place while a wave travels). This was taught clearly during the newly added “Slinky Spring” demonstration (D1 above) showing clearly the vibration (of the hand) which causes the wave to travel though the spring. (C-M and R-M)

D17) Sound travels in electromagnetic waves (Sound travels in longitudinal pressure waves – light travels in electromagnetic waves) A slide on types of waves was added to go with the slide on longitudinal and transverse waves. (C-M and R-C)

6.3 Pilot test on show improvement

As prelude to a full-scale show overhaul and retest, an informal study was performed to try to ascertain whether changes to the show could bring about visible improvements. The study was performed with a group of 24 urban Grade 11 students who happened to be attending a holiday class. Students indicated their willingness to take part by filling in the assent forms on the questionnaire after being fully briefed. They wrote the MCQ pre-test, had a 45 minute show, wrote the MCQ post-test and then spent 45 minutes answering the open-ended survey. The show had not been fully changed, but attention had been taken to address the issues around difficulty D1 (sound travels as a transverse wave) and D16 (vibrations and waves are the same thing). This entailed changing any show slides wrongly depicting sound waves as transverse, and including a demonstration with a slinky spring showing clearly the difference between transverse and longitudinal waves. In addition an animation comparing the two wave types was included in the presentation.

Despite the informal nature of this pilot there were some encouraging results. On D1 (sound travels as a transverse wave), reflected by MCQ 2, 58% of students answering the pre-test wrongly, got it correct in the post-test. On D16 (vibrations and waves are the same thing), reflected by MCQ 1, 48% of students answering the pre-test wrongly, got it correct in the post-test. What was even more encouraging was the answers to Q's 1 – 3 in the open-ended survey, especially the drawings. In our original open-ended study, not one single student (out of 117) had drawn longitudinal waves in answer to any of the drawings requested in these 3 questions. In this pilot study, of 72 drawings, 26 of them (36%) were drawn with longitudinal waves and 30 (42%) with transverse waves. The others drew various other wave-forms. Other results were similar to our main study group.

What was also interesting was the results for OES Q6: "Why does a guitar have a wooden box behind the strings?" Due to an oversight I had forgotten the equipment (small music box) necessary to perform the demonstration attached to this concept. The class results were just 33% of students choosing the correct answer, compared with 65% of the similar urban group in our open-ended survey (who had seen the demonstration). This supports the importance of demonstrations in assisting understanding.

Chapter 7: Redesign of probes (Stage 4)

7.1 Introduction

Chapter 7 addresses the final research question: RQ 4:

> 4) Building on the data obtained from addressing RQ 1, how can the probes be modified so as to be more effective and valid tools for revealing evidence of student learning and any related conceptual and visual difficulties after experiencing *the sound show*?

Section 3.4 provides details of how the probes used in this study were developed and designed through two pilot studies (and in the light of the literature review in Chapter 2 and the synthesis and classification of student difficulties in sound in section 2.6). Chapter 4 provides a comprehensive validation of the probes used in this study which is summarised in table 4.4. The summary section 4.3 concluded that the probes were generally effective with some limitations for the rural students. We cannot make final conclusions about the C-R-M factor if we are using probes which are not valid for all groups. In order to be able to compare results based on an improved show with previous results (chapter 5) it was decided to use the original probes as far as possible and to fine tune these, rather than substantially changing them. It was also felt that some of the issues (like students leaving drawings blank) could be improved by more effective invigilation or the provision of enough time, rather than needing to change the questions themselves. Some suggestions of changes to the probes were discussed in Chapter 4, but these are further elucidated below, with the assistance of a group of interns.

A session was held with 7 student interns at USC at the end of 2015. Student interns are unemployed science graduates who are paid by the government to work for a year in the SC, thereby gaining work experience and skills, while contributing to running the SC. As it was late in the year, the interns had all had considerable experience in the Centre and were ideally suited to comment on the probes, as:

- they had all attended schools similar to the ones surveyed
- they were all isiZulu speaking (so sensitive to language issues)
- they had all both presented and watched science shows similar to “Good Vibrations”
- they had all interacted with thousands of visiting students

The interns were subjected to an identical programme to that used for the student surveys as detailed in section 3.4.3. Before any discussion of the probes they attempted to answer the questions themselves, comprising the MC pre-test and the MC post-test and open-ended survey (OES) after the show. Discussions were informal around a table and were recorded.

7.2 Revision of multiple choice probes

The interns commented (IC) that MCQ were generally good and not confusing:

> IC "In my case, questions were not difficult to understand or answer but I had misconceptions"

This echoes the findings of section 4.1 that good values for difficulties, discrimination indices and distractor analyses show that the MCQ are generally a good measure. Only MCQ 6(b) and 7(c) came out as weak distractors and have been changed in the final MC question sheet (appendix b)

The interns made the following suggestions:

MCQ 4) It was felt that students may identify the word "volume" with the volume control (knob) on a radio (or even with the measure of space), causing confusion. To avoid this, the word "(loudness)" was added to the question after "volume". (This was also done in question 6)

> IC: "For young students in science, a volume can be mistaken as a knob, and not as the quantity of volume"
>
> IC: "Volume could be misunderstood as actual measurements of things in a cup – space"

MCQ 6) Distractor option b ("period") performed poorly in the original study with very few students selecting it in the pre or post-tests (and no urban students). In retrospect "period" is a scientific term which is not familiar to Grade 9 students, and which is not used at all during the show (hence no *more* students chose it in the post-test). Consequently this was discarded and replaced with "wave-speed". This was chosen as a wave-property which would be more familiar to Grade 9's and which could link into other difficulties related to speed (like d4 and d14).

MCQ 7) Distractor option c ("interference") also performed poorly in the original study with very few students selecting it in the pre or post-tests. Again "interference" is a scientific term which is not familiar to Grade 9 students, and which is not used at all during the show (hence again no *more* students chose it in the post-test). Consequently this was also discarded and replaced with "smoothness of the wave-shape". This was chosen as it linked into other difficulties (like d8 and d13) and linked to an option in MCQ 8. It was also used frequently during the show.

MCQ 8) It was felt that the language in the question was difficult for non-English speakers to follow:

IC: “Kids don’t understand punctuation very much . . have to be careful with punctuation. Shorter sentences would be better”

The original question read:

“8) Music is different from noise, because its wave: a) has a smaller amplitude   etc.”

The group agreed to rephrase this as follows:

“8) Why is music different from noise?   a) Its wave has a smaller amplitude etc.”

MCQ 9) Similarly it was felt that question 9 may cause language difficulties:

IC: “In isiZulu it is the tone which distinguishes a question from a statement .. can be confusion when you are reading.”

The original question read:

“9) A *longer* guitar string will: a) vibrate faster, and produce a higher note than a shorter string”

So it was rephrased as follows:

“9) A *longer* guitar string will: a) vibrate faster than a shorter string, and produce a higher note”

Finally the options in MCQ 11 in the post-test were rephrased to mirror those in MCQ 13 of the pretest, to ensure consistency of answers.

7.3 Revision of open-ended survey questions

The interns reported that the OES had been challenging to answer and especially that the time given them was “way too short”. This lack of time may also account for the problem expressed in (4.2.1) that many drawings were not attempted – especially in later questions. Nevertheless they felt the survey was fair and they made suggestions of how the questions could be more clearly stated, and could be more successful in eliciting full student answers. They also suggested clearer instructions to ensure that students answered properly and didn’t just copy the question wording as an answer:

IC: “Brief them –‘please do not copy and paste.’ Maybe print out a set of instructions to whoever is supervising”

Cover page) Interns were happy with the Declaration and Introductions, but felt that the line: "Dear Student we trust that you enjoyed the show …" may pressurize students to respond more favourably to questions and feel obliged to give a good reflection.

> IC "You are already conditioning them, you are forcing them to enjoy your show. You are not giving them freedom to enjoy or not enjoy basically."

Even though this would not be seen until after all the MCQ's were completed, we changed the word "trust" to "hope" to attempt to make this opening line less demanding.

Q1) Interns commented on the fact that the men in the pictures seemed *old*, echoing the findings from student data in (5.2). After viewing a number of possible options for people pictures, the group agreed to go with cartoon "heads" which were free of gender, age or racial biases. The graphics in Q 1 – 3 were all changed to these pictures, as in Fig. 7.1 below. (The group was happy with the cartoon figures used in Q 5 – 9)

> IC "Pictures were old men. Maybe they will think that the old men has a hearing problem. That's why they won't hear. Make cartoons child-friendly."

> IC "Kids will identify with the [new] pictures because they are very animated, emoticons kids know all about it now – will want to relate more to these pictures than to pictures of old men."

> IC" Same pictures for all three drawings important. Consistency. Makes better sense"

Fig 7.1: New graphics for OES questions 1 - 3

Q 1 – 4) The group felt that there wasn't a strong enough link made between the drawings and the written explanations, and that that should be more clearly stated in the question. The original format with all instructions at the top of the page was changed, putting the phrase:

"Show your friend <u>using a drawing</u> how sound travels from A's mouth to B's ear."

above the drawing and then below the drawing and above the lines provided for writing, the phrase:

"Now explain to your friend <u>in words (and with the help of your drawing)</u> how sound travels from A's mouth to B's ear. "

It was felt that this measure, together with more diligent invigilation of this probe could encourage students to attempt the drawings and to link them to their written explanations.This could address the problem identified in (4.2.1) where students provided drawings more as illustrations than explanations.

IC: "You should emphasise that they have link their drawing with the follow-up"

Q2) Figure 5.24 in section 5.2 showed a student drawing where sound goes around the wall and into the room from the back (which was left open on my drawing). To prevent this option, a back wall was added to the room in the new survey.

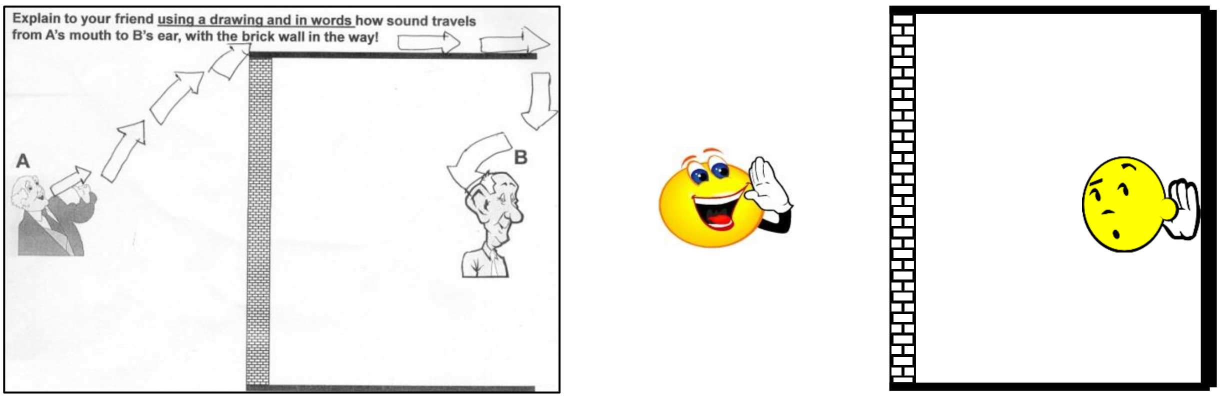


Fig 7.2: Relating to Fig 5.24: Student drawing in response to OES 2 and new drawing

Q 6 – 9) Interns echoed the opinion expressed in section (4.2.2) that students would be unlikely to choose option D ("I don't agree with any of you") as it would require more confidence to then provide their own full explanation.

IC: "They would have to provide their own meaning. That would be hard for them."

Interns also expressed the concern (from their experience) that students may well just repeat the wording in the voice bubbles for these questions, as we found in section 4.2.3.Indeed the student results showed that for these 4 questions, option D was chosen less than 4% of the time, making it a very weak distractor or option. Instead new distractors were written for these questions under option D as follows: (these can be viewed in appendix c)

- Q6: “Why does a guitar have a wooden box behind the strings?” The new option D was: “The box is to make the sounds lower in pitch.” This was added to provide a counterpoint to option A (“higher in pitch”), again reflecting difficulties d8 and d13, both common in the original study and in the literature.
- Q7: “Why does a loud sound hurt your ears?” The new option D was “A loud sound is not music it is noise.” This was added to reflect difficulty d 13 which I found to be common in the original study and in the literature.
- Q8: “Why do men usually have deeper voices than women?” The new option D was “I think it is because men speak louder than women.” This was added to reflect difficulties d8 and d 13, both common in the original study and in the literature.
- Q9: “Can you hear sounds in outer-space (far away from earth)?” The new option D was “You can hear a sound but it will be quieter.” This was added to provide an intermediate option to those in A – C, (yes, no, no) and to reflect difficulties d6 and d9, both common in the original study and in the literature.

Q6) Interns felt that the words used in options A, B and C: “higher, clearer and louder” were too similar, especially for non-English speakers.

> IC: “Clearer, louder and higher sound like the same thing – the words are too similar”

This was born out by section 4.2.3. To make more of a distinction, we changed: “higher” to “higher in pitch” and “clearer” to “more distinct or clear”.

Q8) In the original survey a drawing was not required for Q8 (“Why do men have deeper voices than women?”). It was felt that requesting a drawing may add to what we learn from the student answers, and mean that every question now required a drawing.

Q9) There was much discussion over whether “outer space” would be understood by the students. The wording was retained but the additional explanation (“far away from earth”) added after it.

> IC: “Better if there is a picture as children think there is nothing in outer space. I think they will understand ‘far away from earth’”

All of the above mentioned changes were implemented in the OES and the improved survey is attached as appendix c. The improved survey was validated by the same panel of 4 experts as previously to ensure that each probe does indeed test what it intends to test. As the questions had not essentially changed the same “model answers” used before were kept. A validation of the new probes with student data as done in chapter 4 is shown in the next chapter (Chapter 8). A summary of the changes applied to the probes appears as table 7.1 below:

Table 7.1: Summary of changes made to probes for subsequent survey.

Question	Multiple Choice: Changes
4	"(loudness)" added to make "volume" unambiguous
6	Weak distractor (b) changed from “period” to “wave-speed” "(loudness)" added to make "volume" unambiguous
7	Weak distractor (c) changed from “interference” to “smoothness of the wave-shape”
8	Rewritten so that the question is more clearly stated
9	Rewritten so that the question is more clearly stated
11 post	Options a - d rewritten to match those in Q 13 of the pre-test
Question	**Open-ended Survey: Changes**
Cover	"trust" changed to "hope" so as not to pressurize students
1, 2, 3	Graphics of old men changed to "talking heads" in all 3 questions
1 - 4	Drawings more clearly linked to written explanations
2	Closed off the room to prevent sound entering from behind in drawings
6 - 9	Option D ("I don’t agree with any of you") replaced with suitable distractors
6	Changed: “higher” to “higher in pitch” and “clearer” to “more distinct or clear”.
8	Requested a drawing to assist students to explain (not required originally)
9	"(Far away from earth)" added after "outer space" to avoid confusion.
	General Improvements
	Allow much more time for OES - at least 45 mins
	Better invigilation and encouraging students not to leave blanks
	Clearer instructions to students, especially not to copy question words over.

7.4 Details of the retest: subsequent study (Stage 3: 2nd iteration)

Section 3.3 describes the research design of the original study and includes in table 3.2 a timeline and description of the events and data collection instruments used for the study. In order to allow comparison with the original study, a similar structure was used for the subsequent study. The main difference being that we had new pupils, but drawn from the same three schools as before (Grade 9). A modified probe (subsequent) was used for the survey as described in this chapter (7). A modified show (subsequent) was presented as described in chapter 6. More time was allowed for the open-ended survey after the show. The programme for the three schools was presented over two days instead of one and follow up interviews were conducted with selected pupils at their schools subsequent to their visit to the Science Centre.

On the 31st of May (rural) and 1st of June 2016 (urban and township), the main data-gathering took place at Unizulu Science Centre for the subsequent study – detailed below. The three school groups comprised 42 Grade 9 students each, as in table 7.2 below:

Table 7.2: Biographical details of schools in subsequent study

School type	Situation	Grade	Boys	Girls	Total	Average Age	Black students (isiZulu speaking)	Other students
Urban School	Model C in Richards Bay	9	21	21	42	14.50	17	25
Township School	From Esikhawini	9	5	37	42	14.48	42	0
Rural School	Rural area near Empangeni	9	20	22	42	15.69	42	0
		TOT	**46**	**80**	**126**		**101**	25

The schools were bussed in to the Science Centre to ensure that they all arrived on time and that conditions were identical. Every page of the written survey was numbered with the student's clicker number (CPS system) to ensure that results could be easily compared and correlated. The schedule for each group was as before (outlined in 3.4.3), except that more time was given for the open-ended survey:

Table 7.3: Schedule for each group in the subsequent study

No	Event	Format	Venue	Duration	Data
1	NEW PRE-TEST (Appendix b, Chapter 7)	13, four-part MCQ on PowerPoint, answer with clicker	Auditorium	15 mins	CPS pre-test excel file
2	NEW GOOD VIBRATIONS SHOW (Chapter 6)	Science Show	Auditorium	45 mins	(Informal observations of attitude and involvement)
3	NEW POST-TEST (Appendix b, Chapter 7)	13, four-part MCQ on PowerPoint, answer with clicker	Auditorium	15 mins	CPS post-test excel file
4	NEW OPEN ENDED SURVEY (Appendix c, Chapter 7)	9 Questions requiring: • Explanations • Drawings • Multiple choice (5 / 9)	Classroom	**45 mins**	Biographical Data Questions 1-9 (all on 10 page numbered form)
			TOTAL	2 hours	

I personally administered the pre-test and post-test and presented the show to all three groups, to ensure consistency. I was assisted by staff of the Science Centre, and teachers were present throughout. The show was presented from a script and the same PowerPoint slides used for each group. The show was presented in English and no allowance was made for language difficulties, and no extra explanations given even if it was seen that students were struggling to keep up. The show was precisely timed at 45 minutes and no extra time given to any group. The open-ended survey was conducted by myself as well this time as I wanted to invigilate it more closely to prevent copying and to encourage students to complete the surveys and not to leave anything out. The open-ended survey was written in a comfortable, air-conditioned adjacent classroom, with more time given this time (45 minutes rather than half an hour) to avoid too much rushing.

The cover page of the open-ended survey again gathered basic biographical data (Gender, School, Name and Age) and was signed by students (to indicate full understanding and compliance with the conditions of the survey) and removed. Full ethical clearance for this follow up study was obtained from UKZN (appendix e) and agreement from the relevant schools, and on this occasion from the students' parents in addition. Data collection and processing was done as before (see section 3.5)

In order to delve a little more deeply into the student drawings, mental models and visualizations (C and R-M) especially, follow up interviews were conducted subsequent to this data collection with selected students at their schools. I conducted the interviews myself, recording them on my cellphone, with a small group of students from each school group whose drawings had shown interesting characteristics. No teachers or other staff members were present and the students talked with me in front of their fellow–students. A simple, informal protocol was adopted where I showed the student his/her drawing/s and asked:

- For the student to explain why s/he drew what was there
- For a more detailed explanation of interesting parts of the drawing
- For more detailed explanation on the waves drawn, their shape, why they got bigger or smaller etc.
- Simple extension questions: "If I stood at X, would I hear the sound? Louder or softer?" etc.
- For elucidation with terms which are often confused. E.g. "Loud and soft or high and low."

Student comments relevant to some of the drawings presented were transcribed from the recording and are presented with the drawings in section 9.2.2.

Chapter 8: Validation of Subsequent Probes with new student data (Stage 3: 2nd iteration)

Chapter 4 summarised the development and validation of the instrument used in the original study, and concluded that the original probe (OP) used was valid across all groups, though not necessarily equally valid. In particular, validity was not as clear for the rural group. The subsequent probe (SP) discussed in the previous chapter, will now be discussed in terms of its validity, to address the question raised by RQ 4.

8.1: Multiple choice probe (pre-test and post-test)

To investigate the validity of these new MCQ probes for the intended use and audiences, an item analysis was performed on the subsequent student results for the ten questions, as detailed in section 3.5.2. The item difficulties (p) are shown in table 8.1 below for both the pre-test and the post-test, comparing those obtained originally with those subsequently.

Table 8.1: Item difficulty (p-value) in school groups for pre and post-test comparison, comparing original with subsequent study.

	ORIGINAL			SUBSEQUENT		
	PRE	POST	DIFF	PRE	POST	DIFF
ALL	0.22	0.42	0.20	0.25	0.43	0.18
URBAN	0.30	0.59	0.29	0.31	0.55	0.24
TOWNSHIP	0.18	0.38	0.20	0.23	0.41	0.18
RURAL	0.18	0.27	0.09	0.20	0.33	0.13

When one looks at subsequent **p-values** in table 8.1 for the pre-test for all schools, they were all low with an average of just 0.25. This is exactly the chance value which one would obtain for students just guessing a four-answer MCQ. As the students were naïve and unschooled (in sound and waves), this figure agrees with what one would expect, showing that the results of the questions are a good reflection of the students' *lack of* knowledge in this area. Although slightly higher than the original average of 0.22, it is not very different).

The post-test p-values in our subsequent test are all significantly higher than the pre-test with an improved average of 0.43, again as expected, after the students have received instruction through the show. This suggests that the instrument was valid in measuring the 5 basic concepts in sound (see section 3.4.4), which were covered during the show, as results for all questions were better post-instruction (except for the comparison question 3, as previously discussed).

Again these p-values are low, as in the original study, but it must be kept in mind that the study group again deliberately involved naïve, unschooled students, in order to isolate the effect of the show without background interference from their school. The questions again yielded a good spread of results both across all three groups (standard deviation on post-test = 19.1 %) and within each group, also suggesting good validity. Again, the decrease in average p-value moving from urban to township to rural is consistent with the decreasing resources, facilities and teachers available to each group, (See student context section 3.1.2) indicating that the questions are a good measure of general knowledge of the phenomenon of sound. While these results seem to indicate no improvement from the original study overall, there was measurable improvement for the rural group which will be discussed with the results in section 9.1.4.

A **discrimination index (DI)** was again calculated for each question, comparing the best quarter of students from the post-test with the worst quarter from the pre-test. The results are shown below in table 8.2, including these results from the original study. DI values greater than 0.5 are in bold. The results for combined schools are graphed in figure 8.1.

Table 8.2: Discrimination Index (DI) in groups, comparing original (O) with subsequent (S) study.

MCQ	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	AVER-AGE
ALL SCHOOLS											
Discr Index PREPOST O	**0.57**	**0.77**	0.00	**0.80**	**0.57**	**0.80**	**0.77**	**0.83**	**0.77**	**0.53**	**0.64**
Discr Index PREPOST S	**0.87**	**0.81**	0.16	**0.58**	**0.61**	**0.58**	0.48	**0.77**	**0.74**	**0.61**	**0.62**
URBAN SCHOOL											
Discr Index PREPOST O	**0.7**	**0.6**	0	**0.7**	0.4	**1**	**1**	**1**	**0.9**	**0.9**	**0.72**
Discr Index PREPOST S	**0.9**	**0.9**	0.1	**0.8**	**0.8**	0.4	0.3	**0.8**	**0.9**	**0.7**	**0.66**

TOWNSHIP SCHOOL											
Discr Index PREPOST O	**0.8**	**0.7**	0	0.4	0.3	**1**	**0.9**	**0.7**	**0.8**	0.3	**0.59**
Discr Index PREPOST S	**0.9**	**0.6**	0.3	0.5	**0.7**	**0.7**	**0.6**	0.4	**0.7**	0.4	**0.58**
RURAL SCHOOL											
Discr Index PREPOST O	**0.7**	0.3	0.2	0.3	0.3	-0.1	**0.7**	**0.6**	0.3	0.4	0.37
Discr Index PREPOST S	**0.6**	**0.6**	0	**0.8**	**0.6**	**0.8**	0.4	**0.8**	0.2	0.3	**0.51**

As shown in Figure 8.1, with the exception of the comparison question 3, these values were again almost all above 0.5 (excepting question 7 at 0.48), showing good discriminating potential. Four questions showed excellent discriminating potentials with DI's above 0.71. There were no flawed items with negative DI's and the lowest DI (apart from q 3) was 0.48. Hence I can confidently claim that these 10 questions show good discriminating potential when applied to this combined group from the DI-values, and yielded a good average DI of 0.62, very similar to the original average of 0.64.

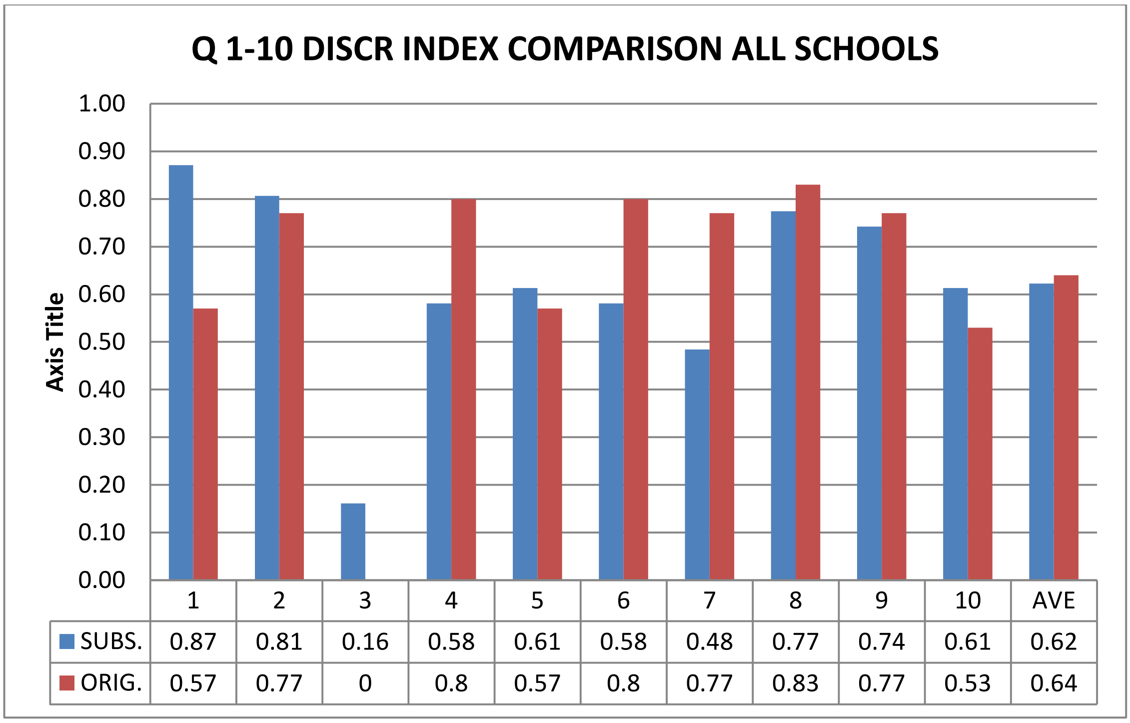


Figure 8.1: Graph comparing Discrimination Indices for all schools, MCQ 1 – 10

In considering individual school groups, the average DI for urban and township schools were both slightly down, but not by very much. This may well be the result of better pre-test scores than originally in both those groups. There is a significant improvement in the DI average for the rural school though which was 0.51 compared with a low 0.37 in the original study. The graph below, Figure 8.2, compares the DI for each question for the rural group:

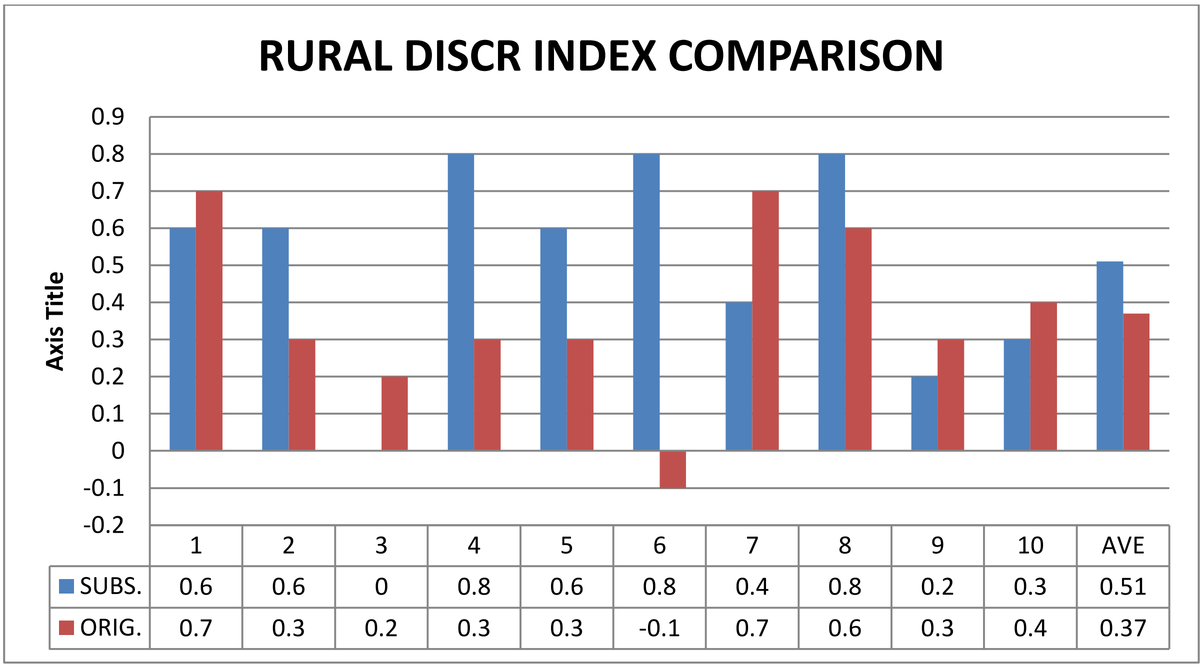
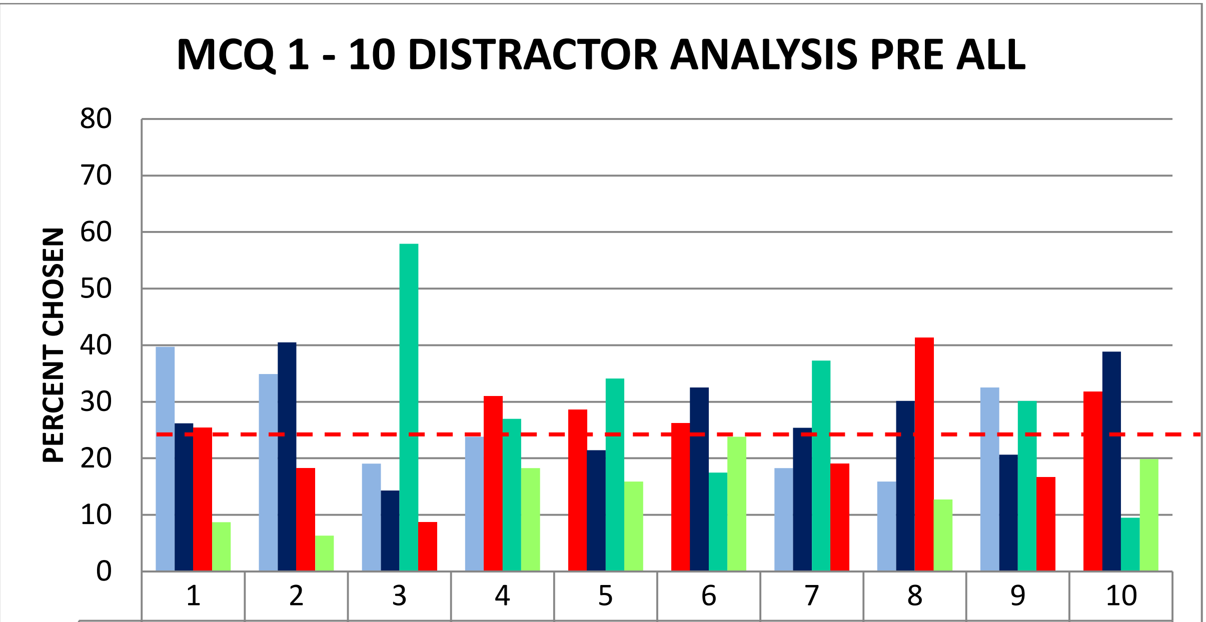


Figure 8.2: Graph comparing Discrimination Indices for rural school.

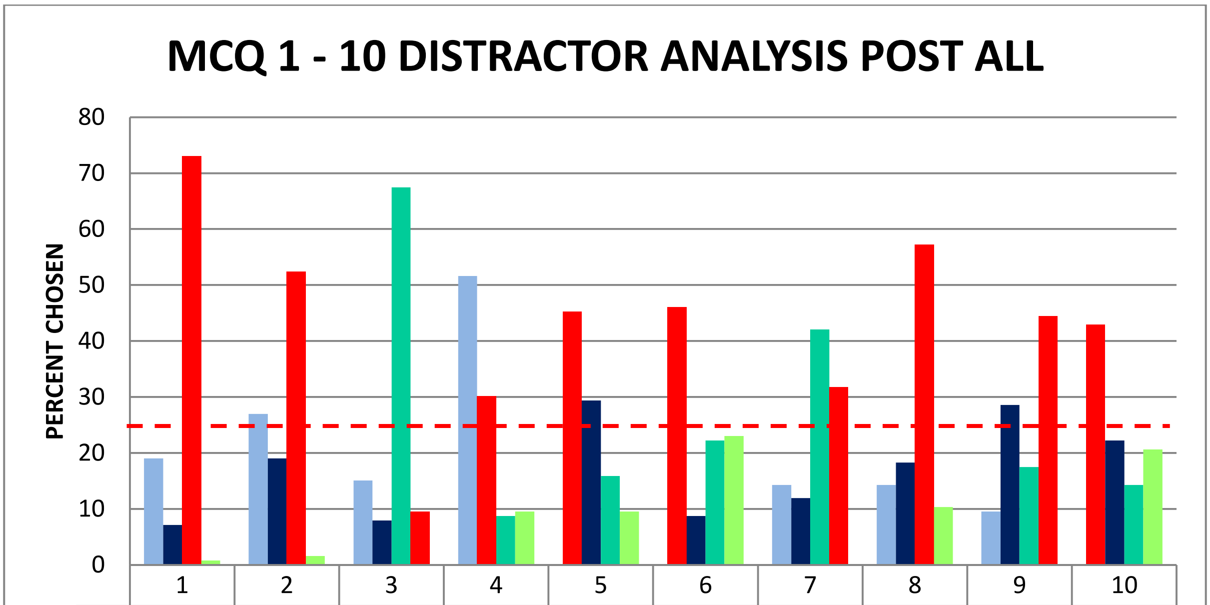
In the original study DI's were generally low, with only 3 questions scoring above 0.5, and one question yielding a negative DI. In the subsequent study, 6 questions and the average obtained DI's above 0.5, showing that the improved subsequent probe had much greater ability to discriminate well between top and bottom students. Hence it appears that the reworked questions greatly assisted the rural groups to understand them, but did not really assist the other two groups. In particular, table 7.1 at the end of section 7.3 indicates that MCQ 4, 6, 8 and 9 were rewritten to try to make them clearer. The graph above shows significantly better DI's for those questions for the rural school (0.8 in each case), with the exception of q 9 where a low DI was still obtained, and the question needs further consideration. (In all these questions the DI's for all schools got worse!)

Distractor Analysis: A distracter analysis was performed on questions 1 – 10 for both the pre-test and the post-test, showing the percentage of students choosing each MC option A to D. The analysis was done for all schools together, rather than individually, as the same MC test needed to be applied to all three groups to allow me to compare and contrast and contrast their performance. In the graphs below (Figures 8.3 and 8.4), the key (or correct option) is coloured red.



	1	2	3	4	5	6	7	8	9	10
A	39.68	34.92	19.05	23.81	28.57	26.19	18.25	15.87	32.54	31.75
B	26.19	40.48	14.29	30.95	21.43	32.54	25.40	30.16	20.63	38.89
C	25.40	18.25	57.94	26.98	34.13	17.46	37.30	41.27	30.16	9.52
D	8.73	6.35	8.73	18.25	15.87	23.81	19.05	12.70	16.67	19.84

Figure 8.3: Graph of Distractor Analysis for all schools, subsequent study, pre-test



	1	2	3	4	5	6	7	8	9	10
A	19.05	26.98	15.08	51.59	45.24	46.03	14.29	14.29	9.52	42.86
B	7.14	19.05	7.94	30.16	29.37	8.73	11.90	18.25	28.57	22.22
C	73.02	52.38	67.46	8.73	15.87	22.22	42.06	57.14	17.46	14.29
D	0.79	1.59	9.52	9.52	9.52	23.02	31.75	10.32	44.44	20.63

Figure 8.4: Graph of Distractor Analysis for all schools, subsequent study, post-test

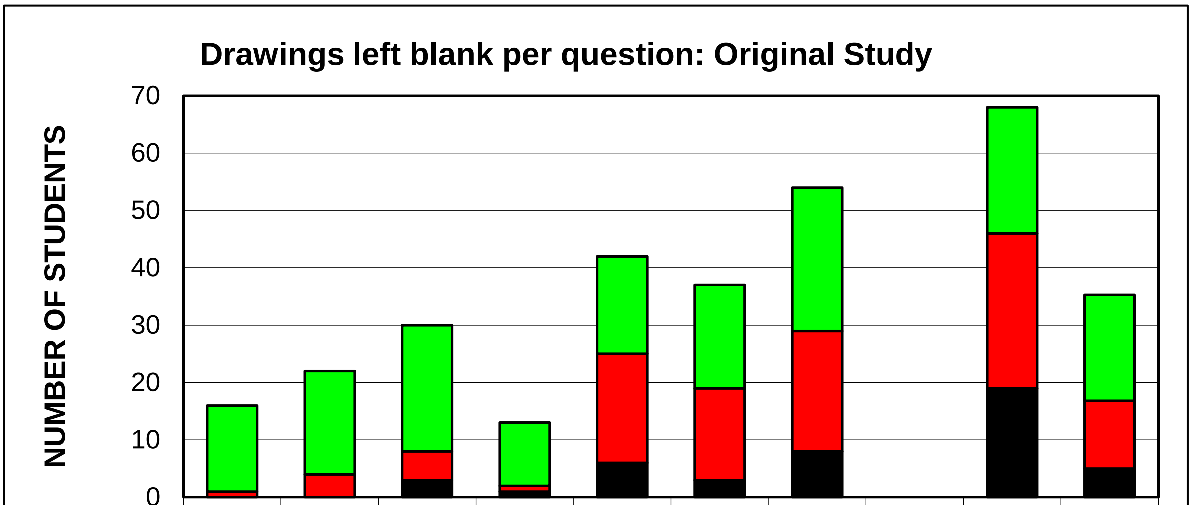
Once again the pre-test distractor analysis graph (Figure 8.3) shows a good spread of choices with no option receiving no choices. Options 1 d and 2 d were not popular and could perhaps warrant further improvement. The new distractors chosen for options 6b and 7c seem to have worked much better than before. Whereas the original options were both the *least* chosen (in the pre- and post-tests) in the original study, the new distractors performed well with 6b being the *most* chosen option in the pre-test, and 7c *most* chosen in both pre-test and post-test. The popularity of this new distracter, 7(c), perhaps points to a new mental image students have of sound where pitch is somehow related to the smoothness of the wave-shape. The post-test distracter analysis graph (Figure 8.4) also showed a good spread over the distracters, but showed that most students had come to the correct selection in almost every question except the comparison question – 3. This meant that the key was the most chosen option, and that percentages choosing the distracters were consequently lower. MCQ 4 showed a trend from the correct answer (b) in the pre-test to the incorrect answer (a) in the post-test: – that the unit of volume is Hertz. This suggests the ongoing difficulty d 8 of students confusing volume and pitch and their units.

In conclusion then, the statistical measures again indicate that my new MC questions are still valid for comparing and contrasting the performances of the different student groups. In particular the improved subsequent probe performed better with the rural group and section 9.1.4 will show that they also obtained better results. Again, the fact that no learning gains were detected with the comparison question(3) – the content of which was *deliberately* not covered in the show – suggests strongly that this probe is valid in investigating whether I can tie the learning achieved to the show. Question 3's results, along with the improvements on the other questions where content *was* covered during the show, convince me that the 10 new MC questions still do actually assess the content in the show, and especially the 5 stated intentions for the show (section 3.2) and not something else, hence showing construct validity.

In addition to this evidence from the data, the improved multiple choice questions were again thoroughly discussed with a panel of 4 experts who confirmed that the questions and answers (underlined in appendix b) are correct and that they do cover the 5 key concepts presented in the show. (This panel comprised two experts in science education and assessment, and two experts in Physics.) In addition, presentations and informal papers covering this work were presented at 2 national Physics conferences and reviewed by colleagues in Physics education who gave further input and made suggestions. The question redesign was informed by interviews with Science Centre interns as discussed in section 7.1.

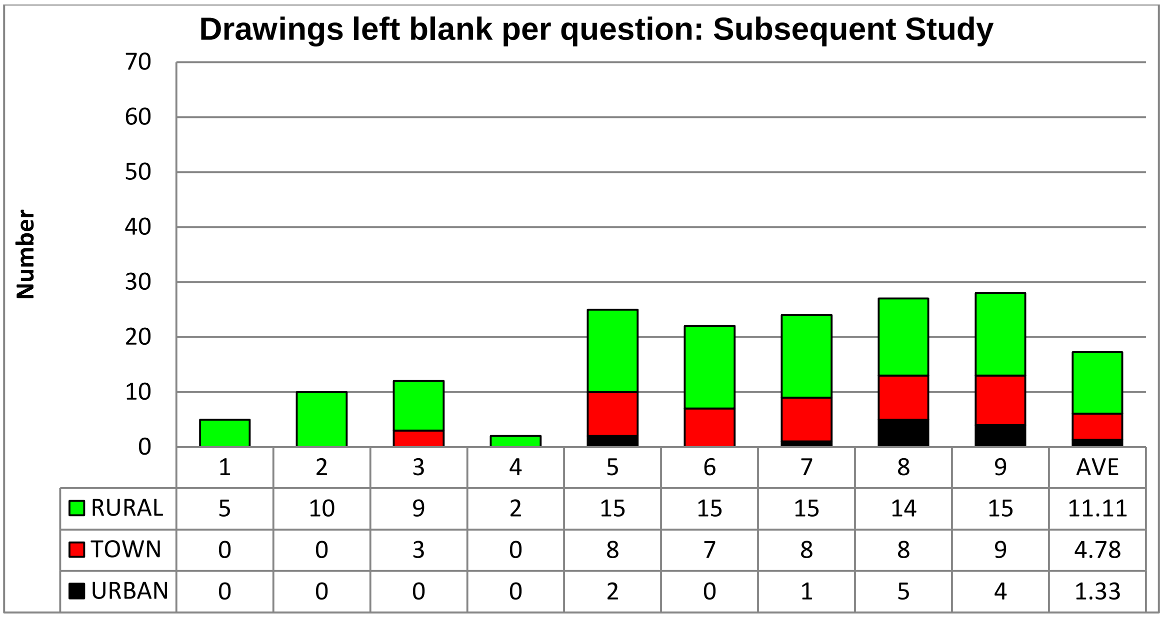
8.2: Open-ended survey: Drawings

As previously reported (4.2.1) the drawings provided very interesting and rich data, but students still clearly struggled to produce them. Nevertheless, it was pleasing that with stricter invigilation and more time to complete the survey, the number of drawings left blank was more than halved (from 282 originally to 128 subsequently – questions 1 – 7 and 9). The graphs below (comparing the original study above, with the subsequent study below) show that despite the improvement, the trends are still the same. Rural students seem least able or willing to provide drawings, and the number of blanks still increases with question number, again showing a “fatigue factor”.



	1	2	3	4	5	6	7	8	9	AVE
RURAL	15	18	22	11	17	18	25	0	22	18.50
TOWN	1	4	5	1	19	16	21	0	27	11.80
URBAN	0	0	3	1	6	3	8	0	19	5.00

Figure 8.5: Graph of drawings left blank in original study



	1	2	3	4	5	6	7	8	9	AVE
RURAL	5	10	9	2	15	15	15	14	15	11.11
TOWN	0	0	3	0	8	7	8	8	9	4.78
URBAN	0	0	0	0	2	0	1	5	4	1.33

Figure 8.6: Graph of drawings left blank in subsequent study

There were no further comments about "old men" after the graphics for Q 1 – 3 were changed but a few comments relating to the new graphics emerged only from rural students (like the selection below), who seemed to feel that A looked happy or loud, while B didn't look good. Whether this affected their answers was not clear. It would appear that any graphic will be subject to some misinterpretation.

(R – Q1) "A sound is loud and B is short because is not laughing at all"

(R – Q1) "Travels smoothly but badly because B doesn't look good"

(R – Q2) "Sound A is travel high because is in the mood but his friend is low"

(R – Q3) "A seem to be happy and louder but B doesn't look so. A is loud and fun."

The "closed off room" in Q 2 did prevent drawings of sound waves entering from behind, but led to some interesting new drawings with waves sneaking through the corners which will be discussed in section 9.1.2. The drawing requested for Q 8 (which was not asked for originally) proved to be a rich source of data and was generally well done, thus adding a useful new dimension to Q 8. A few drawings showed evidence of transfer of understanding from bass-guitar strings to vocal chords, like figure 8.7 below:

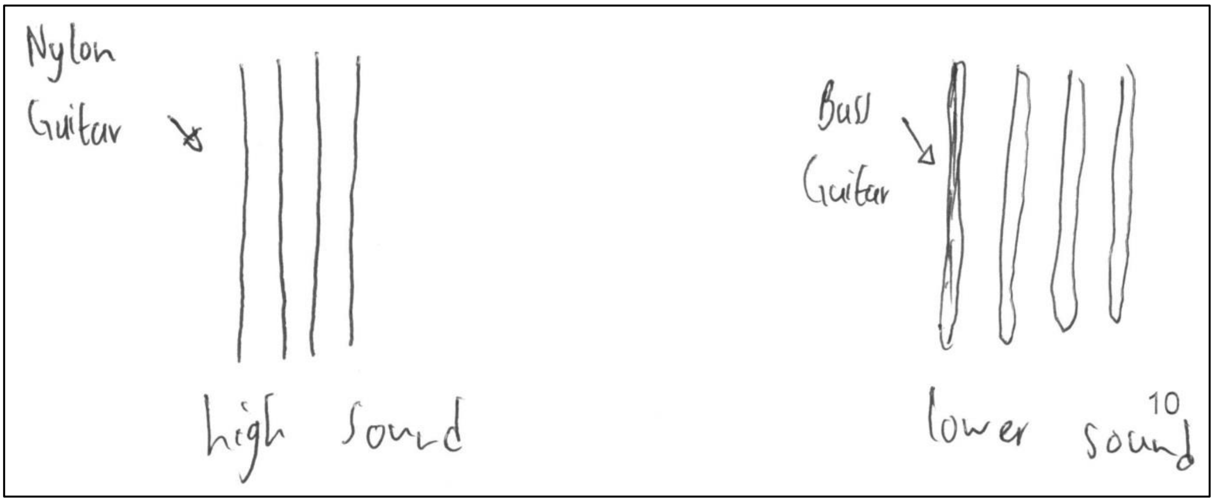


Figure 8.7: Student answer to OES 8

The drawings for Q 9 also showed evidence of students referencing two of the new demonstrations introduced to the show, which also added a pleasing new dimension.

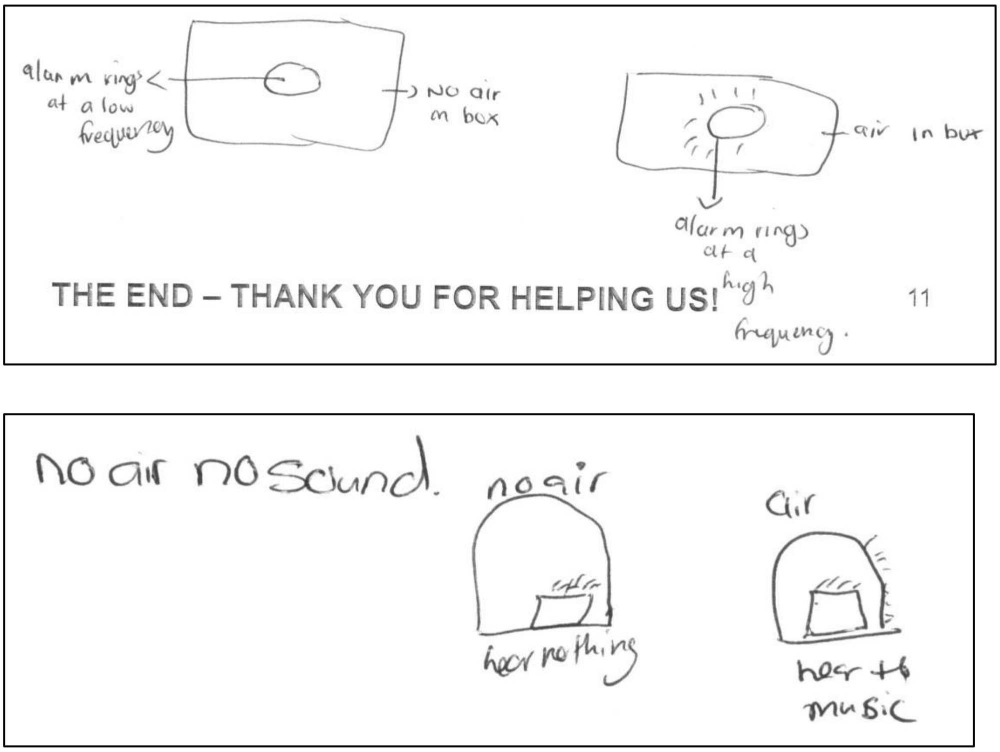


Figure 8.8: Student answers to OES 9 showing reference to demonstrations

Finally some of the early drawings of sound waves (e.g. Q 1 – 3) showed evidence of a "dual-explanation" not seen in the original study – this will be further discussed in section 9.1.2. The results of the subsequent study contain many examples of both sound (providing a science based *explanation*) and unsound drawings (merely an *illustration*) provided by students, validating these drawings probes in terms of their ability to elicit a range of responses from the participants.

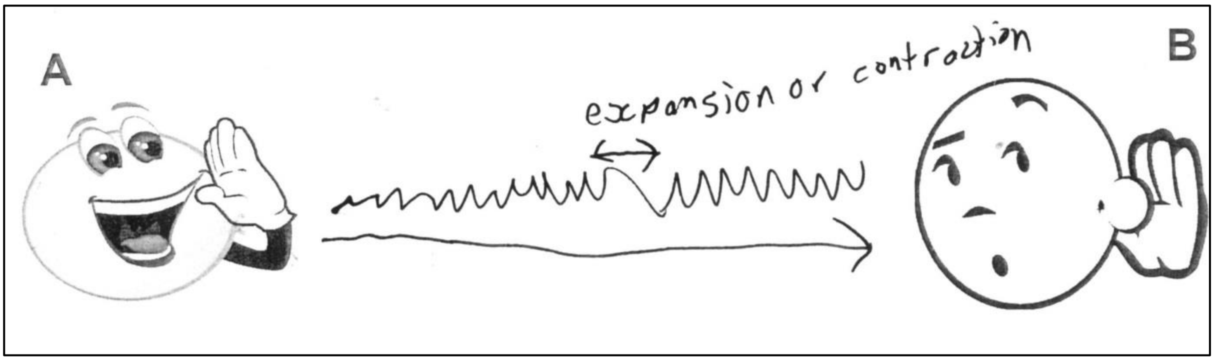


Figure 8.9: Student answer to OES 1 showing dual waves

8.3: Open-ended survey: Choose an option questions. (Q 5 – 9)

Table 7.1 at the end of section 7.3 summarised changes made to questions in this section:

The p-values for Q 5 – 9, graphed below in figure 8.10, are quite similar to those from the original survey, with the exception of Q 8 and 9 where marks were much better. The average p-value of 0.53 is acceptable for a post-test and compares well with the average post-test score for the longer multiple choice questions. The lowest p-value of 0,25 for Q5 correlates with the lowest p-value for MCQ 3 which is a very similar question.

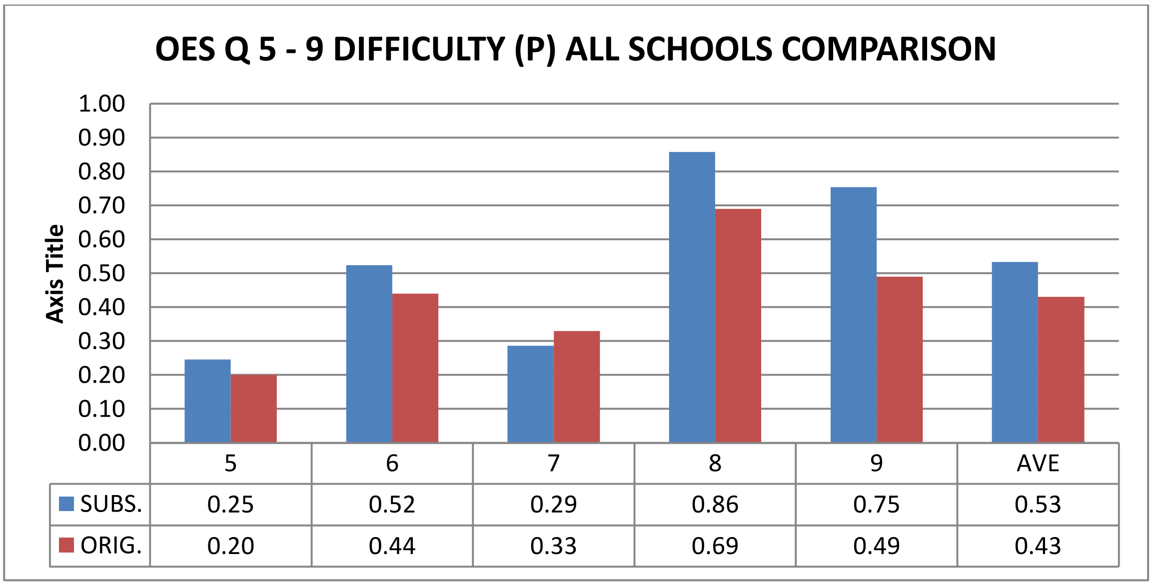
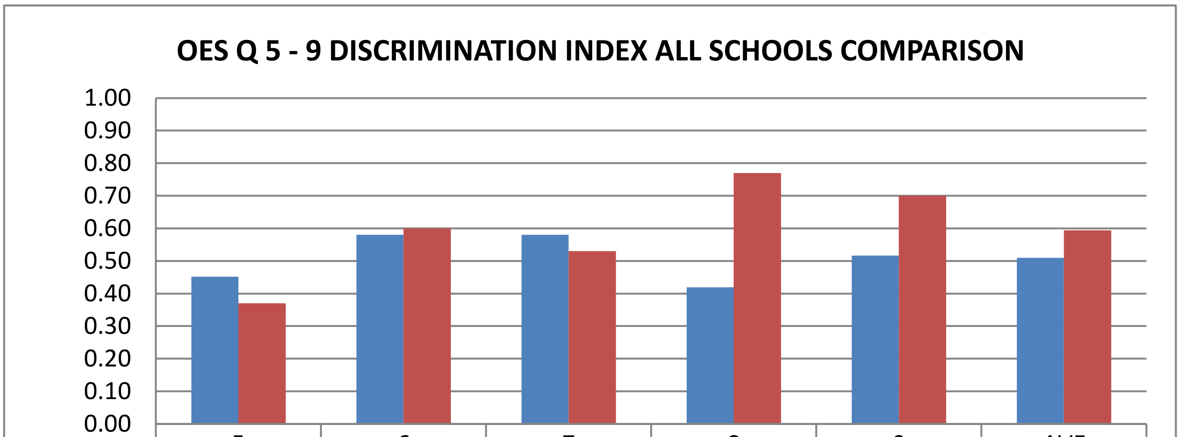


Figure 8.10: Graph of difficulty (p-value) for all schools comparing subsequent and original studies.

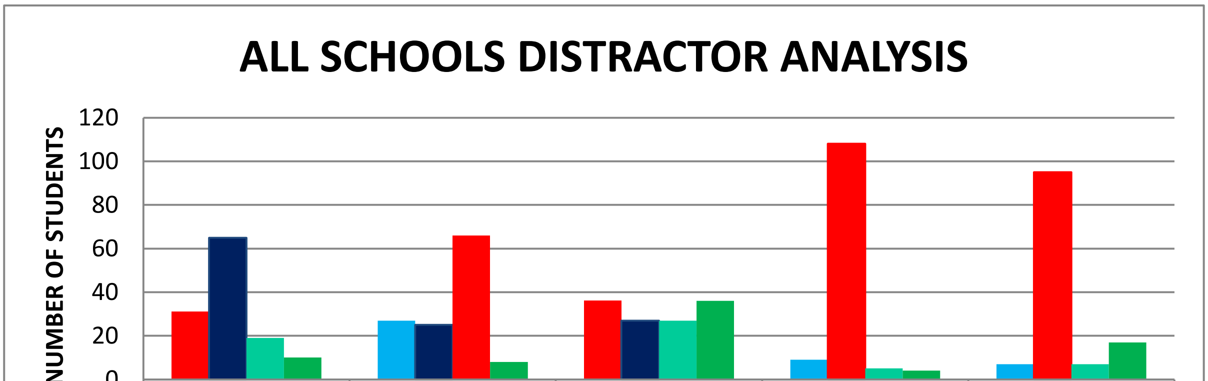
The discrimination index for all schools (Figure 8.11 below) showed acceptable values for all 5 questions (although lower for Q 8 and Q 9) and an acceptable average of 0.51. While lower than previously, the student performance in Q 8 and 9 was much better than before (see Figure 8.10 above) with almost all students getting the right answer, resulting in a lower DI.



	5	6	7	8	9	AVE
SUBS.	0.45	0.58	0.58	0.42	0.52	0.51
ORIG..	0.37	0.60	0.53	0.77	0.70	0.59

Figure 8.11: Graph of DI for all schools comparing subsequent and original studies.

The most difference was seen in the functioning of the new distractors: option D for questions 6 – 9. The distractor analysis graph below shows that these performed well with 16 students on average per question choosing this option as opposed to just 6 before. For questions 7 and 9 option D was the best performing distractor. Having a realistic option for D made all these questions perform better as before they effectively had only 3 options (which students were likely to choose).



	5	6	7	8	9
A	31	27	36	9	7
B	65	25	27	108	95
C	19	66	27	5	7
D	10	8	36	4	17

Figure 8.12: Distractor analysis for OES 5 – 9: all schools (Key is coloured red)

Question 6 had the wording of the options more distinctly defined to avoid confusion. While this did not affect the urban and township students much, the rural students performed significantly better than in the original study (45 % getting the correct answer as opposed to just 16 % before). Again in question 9 where "outer-space" was more clearly explained, the rural students improved (67 % correct as opposed to 43 % before)

In summary the results of this item analysis show that OES Q 5 – 9 still performed well in terms of difficulty and DI, and that the distractor analysis showed that the new distractors (option D in Q 5 – 9) performed much better than before. Furthermore: where wording had been rephrased to make the question clearer, rural students performed better.

8.4: Open-ended survey: Written answers.

As reported in section 4.2.3, written answers were still very disappointing and language was still clearly an issue. Section 7.2 details changes made to the questions in the OES. Most of these were for the purposes of clarity and the questions were mostly unchanged. Consequently student written answers in the subsequent study were not substantially different when compared with those in the original study. There appeared to be less misunderstanding of the questions though and the rewording of these seems to have paid off. Rural students again struggled to express themselves and once one had excluded those who didn't attempt a written answer and those who simply rewrote the question (or the voice bubbles) there was only a handful (generally less than 5) of useful answers.

The questions requiring written answers were again validated by the same panel of 4 experts mentioned previously to ensure that each probe does indeed test what I expect it to probe. The "model answers" against which student answers were measured (shown in appendix j) were essentially unchanged. Again clear examples of sound and unsound student answers were found for each question to confirm that the questions are valid in terms of being able to elicit a range of answers from sound to unsound, but are not shown here as they are not substantially different from those reported in section 4.2.3.

8.5: Summary of Validity

Table 4.4 from the summary of validity in section 4.3 is reproduced below as table 8.3, and a summary is given as to the perceived effectiveness of each probe in the subsequent survey in terms of the objectives below (and in the light of the discussions above), and for each of the school groups. Effectiveness is again indicated using a five point scale, with changes from original to subsequent study highlighted in red, and the previous rating (from table 4.4) in brackets. Some improvements in validity are shown, especially for the rural group, in answer to RQ 4.

The modified multiple choice questions worked well generally, and statistical indicators (especially DI and t-test figures) showed that they performed excellently for urban students and well for both township and rural students. The improvement in validity for the rural group was shown with higher DI indicating better ability to discriminate between strong and weak students and lower t-test scores. The new distractors in questions 6 and 7 performed better than the previous ones.

Open-ended survey drawings were mostly poorly done by all students but it was pleasing to see the number of blank drawings drastically reduced (average blanks per question dropped from 35 to 17 for all schools) with better invigilation. New demonstrations introduced into the show were reflected in some of these drawings. In addition the new graphics ("talking heads") introduced in questions 1 – 3 seemed to cause less problems for students than the old ones. The drawing now requested in answering question 8 revealed some interesting student conceptions.

The "choose-an-answer" questions in the OES performed better on the whole across all groups with the addition of new "option D's" for questions 6 – 9 facilitating a wider spread of answers and section 9.1.2 will show that better results were achieved. While performing well they again caused a problem with the written answers with students mostly just repeating the "voice bubble" options as their answers. Written explanations again did not perform well on the whole and were a reflection of students' weak English language abilities with rural students' answers being so weak as to be almost unusable.

On the whole it can be seen that the new probes did not make a marked difference for the urban and township groups (although improved invigilation yielded more data), but appeared to be more valid for the rural group in every part except the written answers. As they were the group that struggled the most with the probes in the original survey, I feel that targeted progress has been made here and that the probes can now be confidently used across all three groups (although written answers still present a major problem for township and rural groups because of poor English ability)

Table 8.3: Summary of validity: subsequent survey (previous ratings from table 4.4 in brackets):

PROBE IN SUBSEQUENT STUDY	1. Knowledge?	2. Purpose?	3. Valid for what kind of student?
Pre-test MCQ 1 – 10	Baseline prior knowledge of 5 key concepts in sound (see 3.4.4) GOOD	Establish a baseline or starting point for designing show or teaching sound GOOD	U – EXCELLENT T – GOOD R – GOOD (AVERAGE) (Better DI)
Post-test MCQ 1 – 10	Assess learning during show in terms of increase in knowledge of 5 key concepts in sound EXCELLENT	Measure what has been learnt during the show. EXCELLENT	U – EXCELLENT T – GOOD R – GOOD (AVERAGE) (better rural performance, DI and t-test figures)
Open-ended Survey Q 1 – 9 Drawings (Post-test only)	Student visualization of concepts in sound and waves. GOOD	Try to see sound concepts through the eyes of students AVERAGE (Students still very weak in drawings)	U – GOOD (AVERAGE) T – AVERAGE (LIMITED) R – LIMITED (VERY LIMITED) (more willing to attempt drawings, more demonstrations reflected)
Open-ended Survey Q 5 - 9 Choose an Answer (Post-test only)	Student conceptions in sound and waves. GOOD	Probe more deeply into student conceptions by assisting them with suggested answers. EXCELLENT (GOOD) (better with new distractors)	U – EXCELLENT (GOOD) T – GOOD R –GOOD (AVERAGE) (new distractors BUT still affects written answers adversely)
Open-ended Survey Q 1 – 9 Written Explanations (Post-test only)	Student conceptions in sound and waves. AVERAGE	Probe more deeply into student conceptions AVERAGE	U – GOOD T – LIMITED R – VERY LIMITED (due to English ability)

After a second round of testing, and despite improvements to the OES reported in the chapter, I am now even more convinced that while the OES yields interesting results, it is ultimately better suited to a classroom situation where instruction is more systematic and students have time to develop their own concepts and visualizations, than to one forty-five minute show. It is too long and detailed to be performed as both a pre- and post-test, so it was difficult to use it as a tool to measure learning in the show. It is best suited to looking at student preconceptions and should perhaps in future be administered before the show on a naïve, unschooled group, to provide an understanding of student's base knowledge on which to build the show.

With regard to my struggles to make sense of written answers: diSessa (1988), speaking of college students, stated that: "students come to Physics classes with no theory at all, but instead are used to dealing with the world on a catch-as-catch-can basis." He further asserted that far from having theories, students have: "a fragmented collection of ideas, loosely connected and reinforcing, having none of the commitment or systematicity that one attributes to theories". Cooke and Breedin (1994) concluded from their study of students that the students lacked explanatory coherence, that they have difficulty providing complete explanations and that they would often offer descriptions rather than explanations. All this was especially evident with my Grade 9 students and exacerbated by their extremely weak English language ability, casting doubt on the efficacy of questions requiring the production of written explanations from scratch.

Cooke and Breedin (1994) also found in their study of experienced (employed in Physics) and inexperienced candidates that selection questions worked better than production questions for the inexperienced group (but both had the same effect for the experienced group). As my students were completely inexperienced (or naïve) this lends credence to my belief that selection questions will ultimately work most effectively for these groups. Consequently, recommendations are made in the final section, 11.6, on a manageable, practical testing regime suitable for the science centre situation, for which multiple choice questions appear to be more appropriate.

Chapter 9: Results and discussion: subsequent study (Stage 3: 2nd iteration)

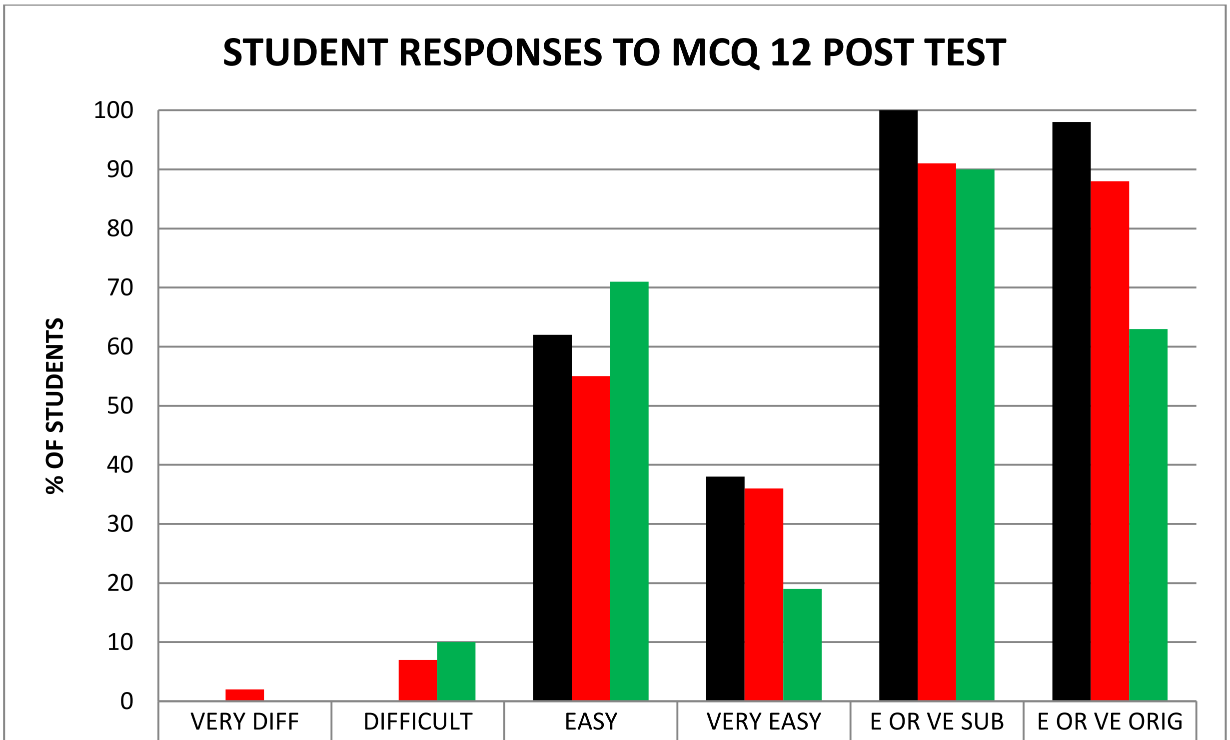
9.1 General attitudes of the student groups (RQ 2a)

As in section 5.1, responses to the attitudinal questions (MCQ 11 – 13 pre and post-test) are presented for the new groups in the subsequent study. This section will be just a brief summary of these results: - full details, statistics and graphs are given in appendix o. These subsequent results were mostly very similar to those from the original study, with the exception of question 12 in the post-test (which reflects more directly on understanding of the new show itself). This similarity suggests that we can disregard differences in attitude between the subsequent and the original groups as a major factor affecting results.

Results of pre-test q 11 reflected that of the subsequent cohort of students, it was the urban group who had visited the Science Centre the most (25 out of 42 students visiting once or more times) with township and rural having visited less (16 and 14 students respectively visiting once or more). Pre-test q 12 reflected that again most students would take *some* science in Grade 10. There is again a strong agreement between all three schools, indicating that we can rule this out as an extra variable in evaluating their responses. As before, the highest dropout rate from Physical Science (11 out of 42) came from the urban school, and was only 2 and 3 students for the other two schools.

Pre-test question 13 reflected that it is again surprising that those at the best school find science to be less interesting – perhaps the reason for the higher dropout rate reflected above. Only half of the urban students (23 out of 42) found science at school to be enjoyable or very enjoyable, but more than 75 % of students at other schools did (32 for township and 33 for rural). This contrasts with enjoyment of the science show (post-test question 11) where almost 90 % of students in all three groups reported that the show was enjoyable or very enjoyable, again indicating that we can disregard enjoyment of the show as a cause of different results between the groups.

Post-test q 12: “How understandable was today’s show?” was answered quite differently in the subsequent study, with far more rural students stating that the show was easy or very easy to understand than in the original study (38 out of 42, or 90%, compared with 63 % before). Urban and township scores were almost identical to those in the original study). This (along with other data presented in the next sections) would seem to suggest that the improved show was easier for the rural group to understand. This is reflected in Figure 9.1 below. The second column from the right sums answers for “easy” or “very easy” in the subsequent study, and the right hand column shows this sum from the original study.



	VERY DIFF	DIFFICULT	EASY	VERY EASY	E OR VE SUB	E OR VE ORIG
URB	0	0	62	38	100	98
TOWN	2	7	55	36	91	88
RUR	0	10	71	19	90	63

Figure 9.1: Graph of student responses to MCQ 12 in the post-test, subsequent study

Answers to question 13 in the post-test (novelty value of the show) indicate that it was probably again misunderstood by the rural school, and is not a good reflection of the novelty value of the show. This question needs careful reworking if it is to be useful in future.

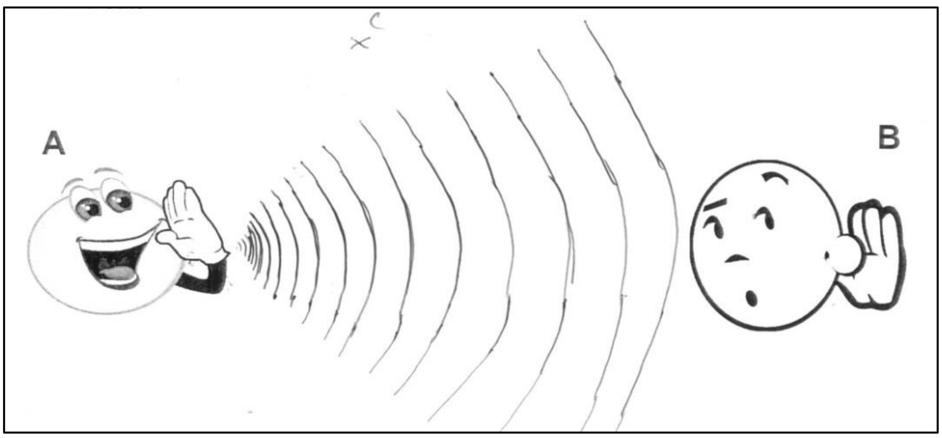
9.2 Students' conceptions and mental models with sound (RQ 2b)

As revealed in drawings

The subsequent *questions* requesting drawings as answers were essentially unchanged from the original study, with the exception of adding a drawing to question 8, where none was requested previously. Section 8.2 reported a marked increase in the number of drawings attempted, which provided more data to analyse – especially for the rural group. Drawings were classified using a similar coding as in the original study (with codes for new drawings added) to allow for comparisons. As these drawings were extensively analysed in section 5.2, I shall comment here mostly on changes evident in the subsequent study.

OES Questions 1 – 3 asked students to draw the sound waves travelling between two people in different situations. There were many changes to the show (see sections 8.1 and 8.2) to try to deal with our main student difficulty (d1) of sound travelling in transverse rather than longitudinal waves. While table 9.2 shows measurable improvement in answering MCQ 2 correctly, the same improvement was not evident in drawings. There were a few students who drew curved longitudinal waves in answer to these questions (7 or 8 in each question) while there were none in the original study. Still the vast majority (about 60 % in each question) drew transverse (or triangular transverse) waves. The promising improvement in this area reported from the pilot study in section 6.3 (36 % drawing longitudinal waves) was perhaps due to the fact that those students were Grade 11's and had developed a scientific sophistication allowing them to better distinguish between representation and reality. It is perhaps questionable in the light of this whether we should reasonably expect a major shift in visualizing sound from naïve students after such a short interaction with them. Even once the symbolism in the ER (C-M) had been addressed and diagrams corrected, it appears the students prior knowledge and mental models (C) are still what comes out in their answers.

Some examples of sound student drawings are shown below in figure 9.2 for OES questions 1 – 3: this concentric circle pattern was not evident in drawings from the original test so some students appear to have managed to reason correctly with the ER (R-C).



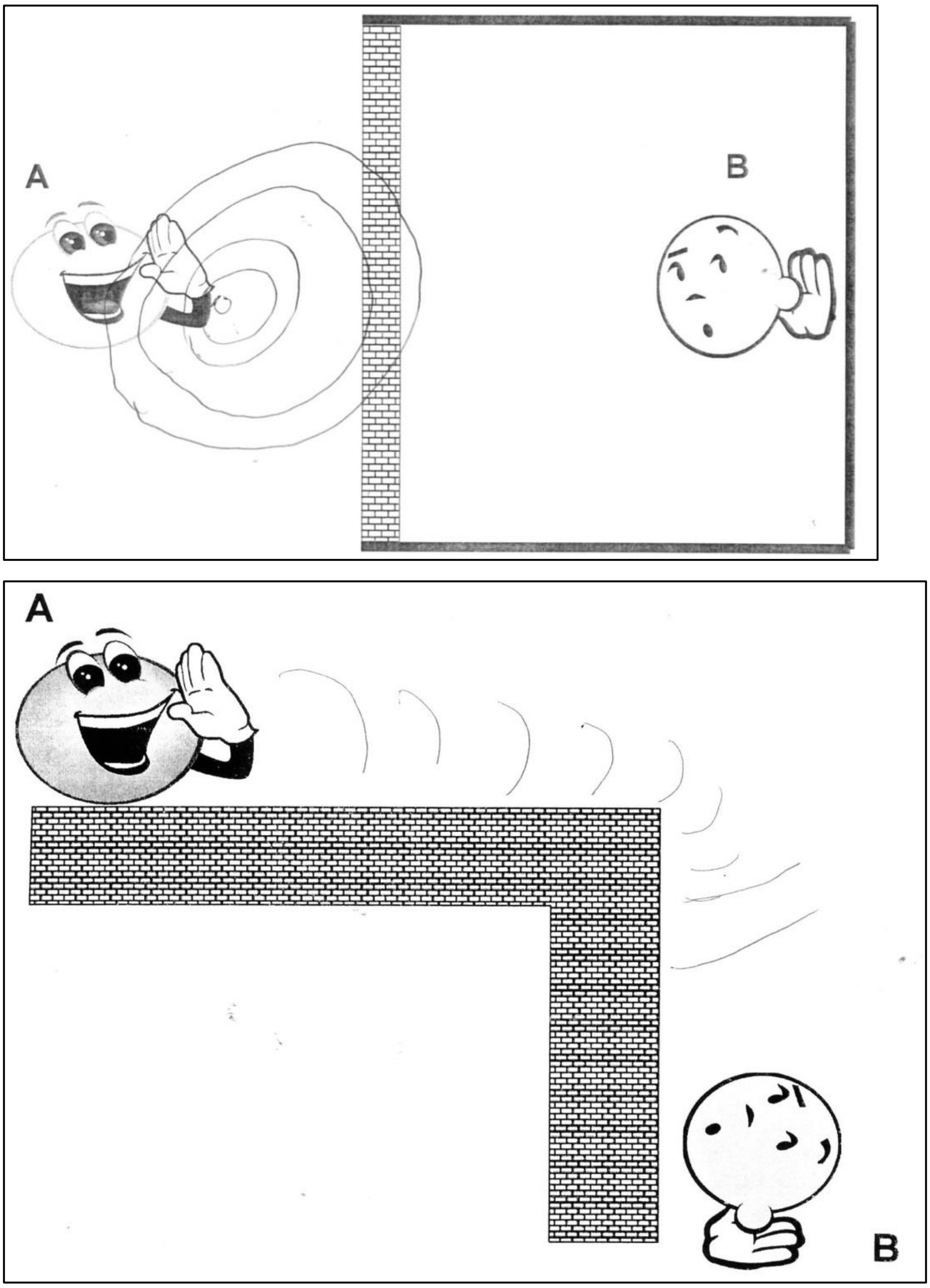


Figure 9.2: Student drawings (U, R, Z) in answer to OES 1 – 3.

It was interesting to hear more about the first drawing in a follow up interview: the student explained that rather than showing the sound spreading out, the “bigger lines show the sound is lower – can’t hear. Short means loud but long means low.” Again this ER (show) seems to be just too short and shallow to deal with deeply entrenched mental models and prior conceptions (C).Nuances like this are not clear without follow up interviews though. The student producing the drawing immediately above explained that the sound “goes well [around the corner] – nothing is in the way” rather than describing it as spreading out or spreading around the corner. Again the attempt to describe a wave spreading out in the ER has not transferred to the student (R-M).

An interesting new phenomenon in the subsequent study was the appearance (from 2 students) of sound waves resembling travelling concentric circles like the one in figure 9.3 below.

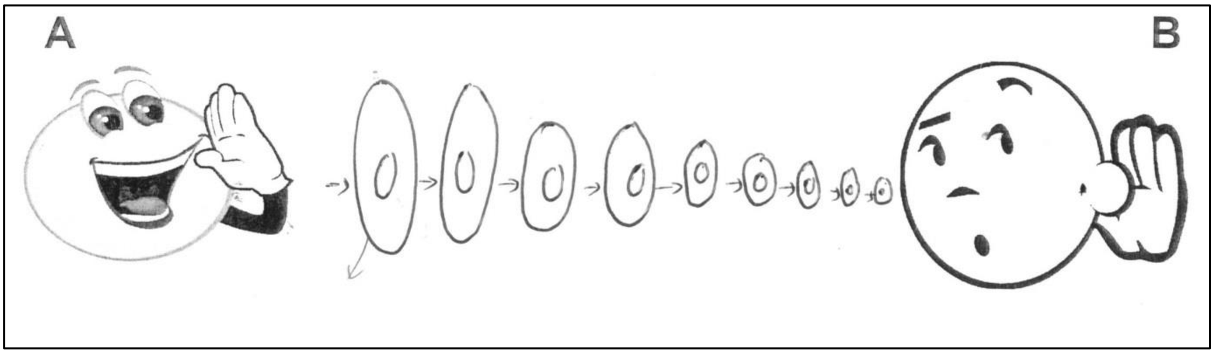


Figure 9.3: Student drawing (U) in answer to OES 1.

One of the changes to the show was the introduction of a new slide comparing sound waves to ripples on water after a drop falls. The student who produced this drawing confirmed in an interview that she had obtained the idea from here (“from the ripple in the water”) and that she would not have drawn this before the show. This student described the diminishing circles as an indication of decreasing volume because it is “louder at A than at B.” It is an interesting example of Linder’s (1992) finding that sound is seen as an “entity” carried by the molecules. Again the student has not managed to reason within the ER and understand the notion of a wave spreading out concentrically from a source (R-M). Also of interest was the appearance of “double drawings” where students provided two different pictures of sound, like in figure 9.4 below.

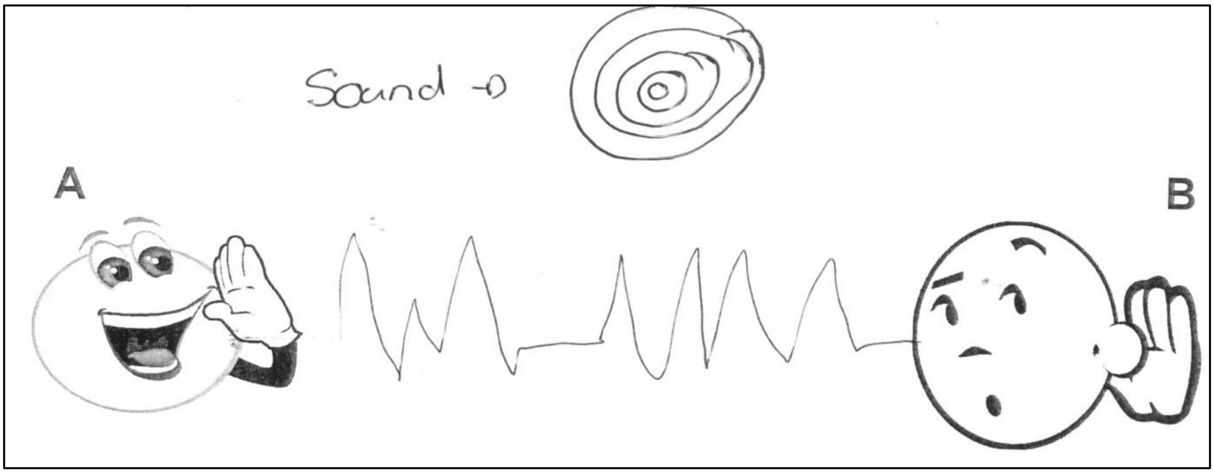


Figure 9.4: Student drawing (U) in answer to OES 1.

This student said that she made 2 drawings to give “another opinion”, and that the wave at the bottom was “like sound, loud and soft” and the circles at the top “shows it goes everywhere, anyone can hear what you are saying.” Again a problem at the R-M level with the student not quite sure (from the ER) of how to represent waves. This perhaps also suggests a blended view of sound as an entity having wave properties, such as that proposed by Hrepic et al. (2010).

The drawing for **OES question 2** had been changed to prevent sound going around the back of the room, but 2 students now indicated sound travelling through the corners of the room, as in figure 9.5 below, reinforcing difficulty d5 that sound needs small holes to escape. This student said that even if there were no holes the person [B] “would hear the sound but softer.” On being asked how the sound finds the hole, she stated “it looks for an exit” and there are other sound waves but the one that finds the exit “is the clever one.”

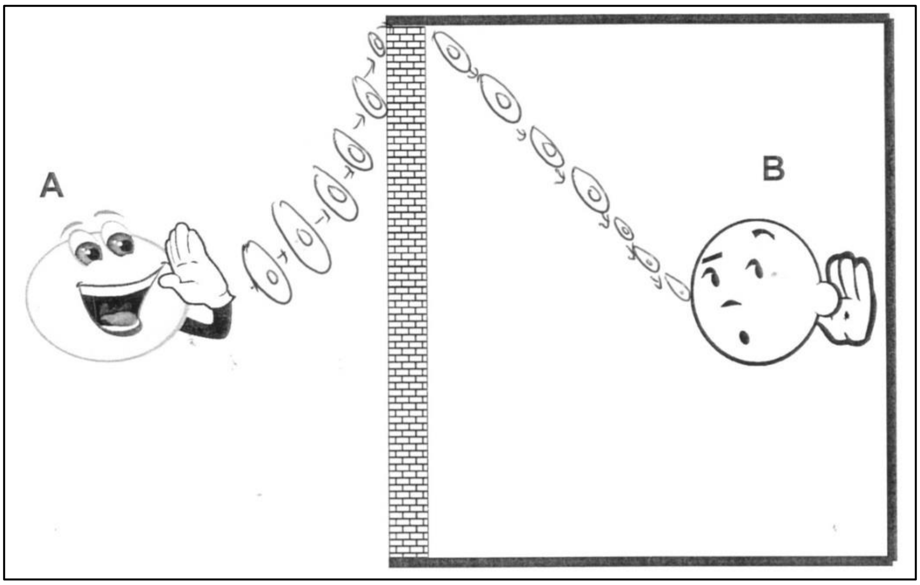


Figure 9.5: Student drawing (U) in answer to OES 2.

What was pleasing to see was that 29 % of students showed the sound as lower (or smaller) after the wall compared with just 8 % in the previous study, as in figure 9.6 below. The introduction of the demonstration with a cellphone ringing in a closed lunchbox may be responsible for this improvement. Most students guessed that the cellphone would be inaudible when the lunchbox was closed and were surprised to finds they could still hear it, but much quieter.

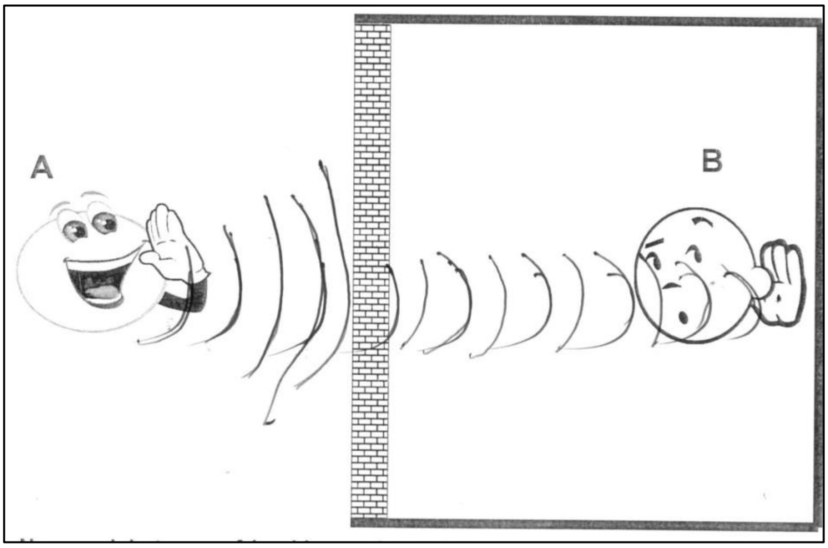


Figure 9.6: Student drawing (T) in answer to OES 2.

OES Question 4 requesting a drawing of the difference between noise and sound waves was well answered on the whole with 66 % of students (compared with 56 % in the original study) providing a correct drawing of the two waves. As this flows directly from a demonstration in the show it is pleasing that students appeared able to reason with the ER here (R-M), unlike in the next question. Drawings for **OES question 5** on the speed of sound, which was deliberately not covered in the show as a comparison question, were consequently poor as it was not covered in the ER. Very few (5 %) showed any particles or any microscopic considerations. What was new was the emergence of pictures relating to two of the demonstrations added to the show (cellphone in a lunchbox and the bell in a vacuum video) like the one in figure 9.7 below:

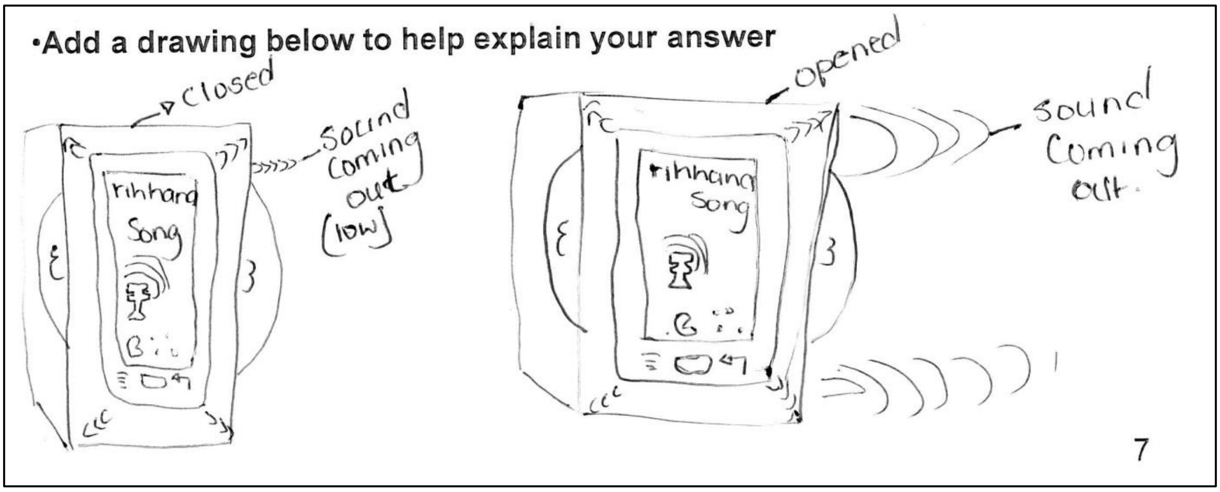


Figure 9.7: Student drawing (T) in answer to OES 5.

Again it appeared that our water drop analogy caused some confusion here as is evident in figure 9.8 below. Water waves are taken as sound travelling through a liquid. There was definite evidence of Linder's "sound resistance factor" (Linder 1993) in the following interview about this picture: "In liquid, if we are putting water on something big like a bowl can see waves spreading out so that you can even hear it. Sometimes that there is a sound. Gas is the same like air . . . sound can travel through fast without anything that can make it change. In solids air can travel inside the solid but not that clear. . . solid can take vibration easily." It seems that the surface water waves (in our analogy) are seen as the same as sound waves travelling through water. I should perhaps have heeded Linder's (1992) warning that the "water-wave analogy" can be profoundly misleading! In future I think it would be worth removing this analogy from the ER as it is questionable at a C-M level and appears to hinder students as much as to help them to reason with the ER (R-M).

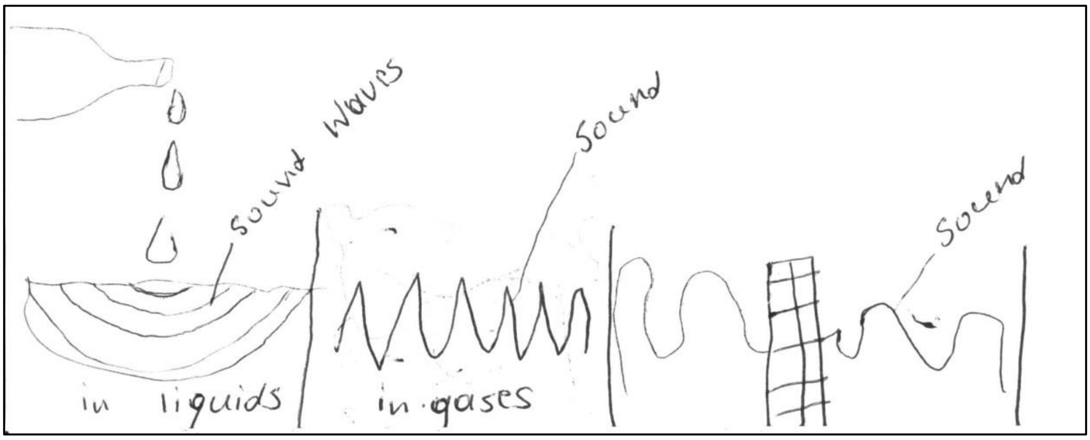


Figure 9.8: Student drawing (Z) in answer to OES 5.

In **OES question 6** (figure 9.9 below) similar percentages of students (10% in both original and subsequent studies) drew the demonstration placing the music box on the guitar in answer to the question about the sound box of a guitar.

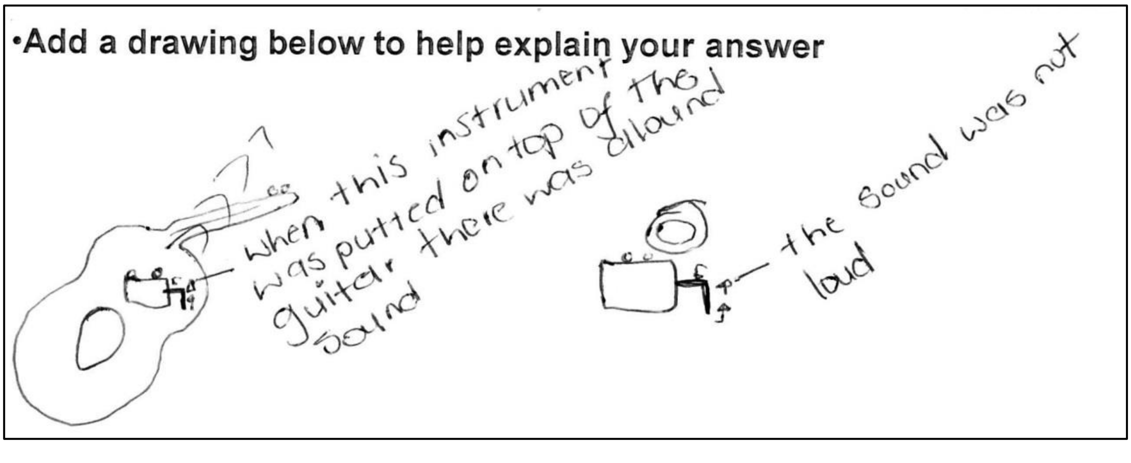


Figure 9.9: Student drawing (T) in answer to OES 6.

In **OES question 7** (about loud sound hurting our ears) one drawing (figure 9.10 below) reflected a slide in the show concerning a "teenage repellent" device, but incorrectly confusing volume with pitch and noise. Again a question not directly dealt with in the ER.

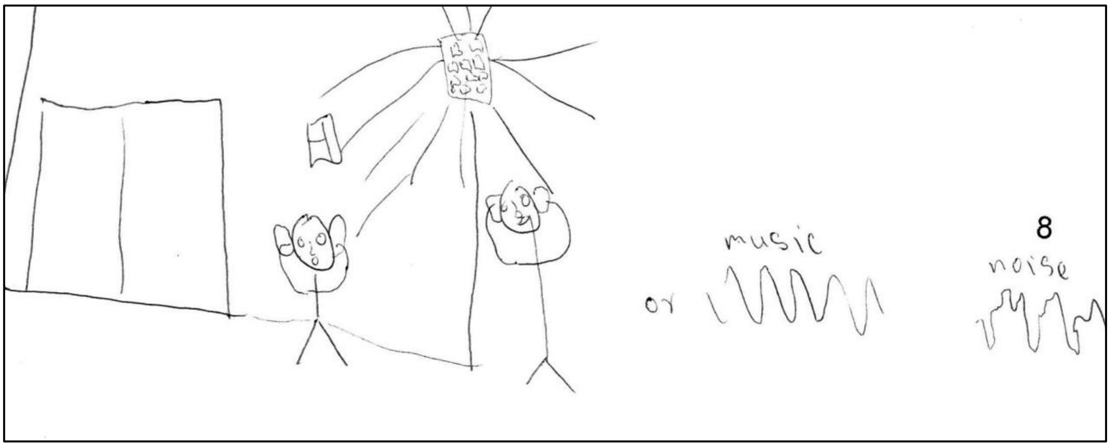


Figure 9.10: Student drawing (Z) in answer to OES 7.

The drawings for **OES question 8** were not requested in the original study, but added some interesting insights into students' thinking. 10 % of these showed vocal cords (though mostly from the urban school) and indicated how these change from boys to men, as in figure 9.11 below. These were few in number but were pleasing to see as they were not directly shown in the ER thus required students to transfer knowledge (R-C) as illustrated in the "guitar strings" drawing (figure 8.7).

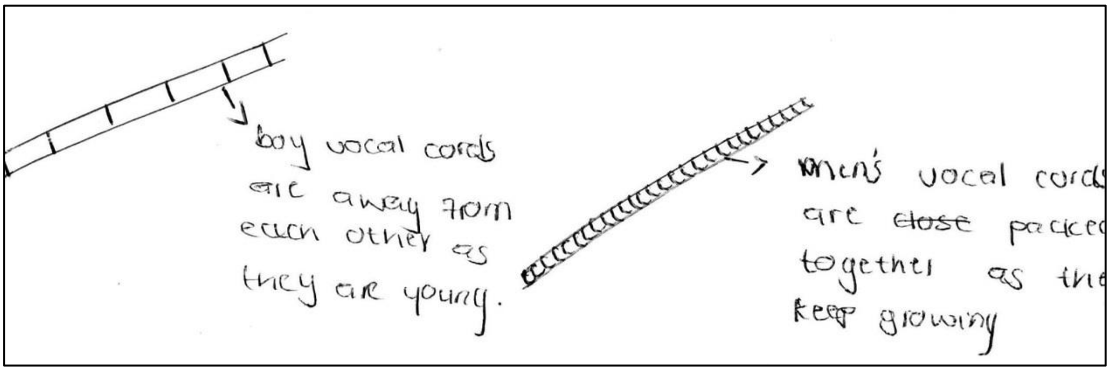


Figure 9.11: Student drawing (R) in answer to OES 8.

Another addition to the show was a graph showing the change in frequency in male and female voices with age. Almost a quarter of students (23% across all school groups) draw some version of this graph, a third of those mostly correctly. The graph in the show is shown in figure 9.12 below alongside a typical student drawing. While it is pleasing that students showed recall of what they had learnt in the show(R-C), the graph does not really explain *why* men have deeper voices than women, but merely gives age details. Hence some aspects of R-M were still missing.

Figure 9.12: Student drawing (T) in answer to OES 8 and graph from the show

Finally **OES question 9** once again saw mostly *illustrations* of a space scene (rather than *explanations*) but 5 urban students correctly drew the bell jar vacuum experiment (a video of which was shown during the show) as in figure 9.13 below and 6 township students drew the lunch box experiment, which is less applicable here. The student producing the drawing below stated (after struggling a while to remember) that it was based on the video of the bell-jar experiment in the show. She explained her drawing that: "There is nothing to carry the sound like air." It was pleasing to see specific demonstrations used to answer related questions and evidence of transfer of knowledge (R-C) from the video in the show to a real-life situation.

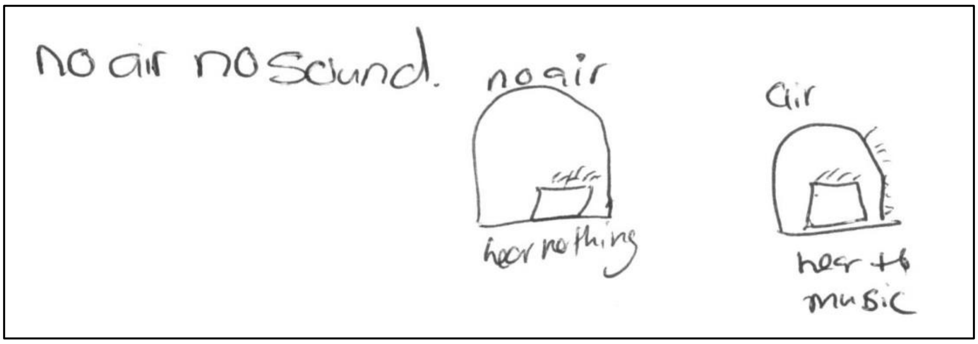


Figure 9.13: Student drawing (U) in answer to OES question 9.

The scope of this thesis does not allow all the drawing examples to be shown for the subsequent test set. While these drawings again revealed fascinating insights on the students' ability to reason with the ER (R-M), they were not substantially different from the original drawings, despite changes to the show and fine-tuning of the probes (notwithstanding the few items of interest mentioned above). I believe this exercise was too limited in scope and time to allow real progress to be made by students in their visual grasp of all facets of the ER (R-M). While we should always make sure that the visualizations and models presented in a show (M) are as helpful as possible, I think deeper, more prolonged interaction with these is necessary in order to see a real change from students at the R-M level.

As revealed in written answers

As in the preceding section, I shall report only on aspects of the written answers in the subsequent survey which were substantially different from those in the original survey. Written answers were still very poor and lack of English language ability appeared to be the major factor here. Once again the "voice bubble" options in OES questions 5 – 9 were a hindrance to this section as many pupils simply repeated their wording as their answer. Rephrasing some of the questions for clarity (section 7.3) seemed to help with the "choose-an-answer" questions (see below) and to some extent with the drawing questions, but not with written answers. Even when the question is clearly stated, students struggle to express their answers coherently.

While the drawings in response to OES questions 1 – 3 revealed a few interesting new ideas, the written answers did not, and very few really explained what was happening satisfactorily. A large proportion of students simply said that A's voice must be louder (or higher) for B to hear him but did not explain the mechanism. Answers were very confused as this example from a rural student shows: starting off promisingly with mention of longitudinal waves, then getting mixed up:

> (R) "A: sound travel using a longitude wave. B hear the sound but not in a high wave, but in a short wave because there are up and down the sound. Sometimes move to the left and to the right using transverse sound devise"

OES question 4 was answered better but did not reveal any new insights in the written answers.
OES question 5 (not covered during the show) was poorly answered but written answers did reveal again the confusion with water waves mentioned in the previous section. Two rural students who felt that sound travelled fastest through liquid offered the following explanations:

> (R) "When you drop a drop on water its waves spread amongst the water - sound travel great in water because its liquid anything can travel through water."
>
> (R) "When you hear a drop of water falling in a dam you can hear it far way."

OES question 6 was similarly poorly answered but the written answers (as well as the drawings) did make mention of the new cellphone in a lunchbox demonstration:

> (R) "Cellphone in box will hear the sound lower"

The students' responses to OES question 7 showed obvious confusion between music, noise, volume and pitch, while some answers revealed the tendency to give "life-experience" explanations, like this rural example (which, though charming, does not really shed scientific insight on the question):

> (R) "Music is something that is needed in your body not noise. Noise closes people's ears but music make people dance or move muscles. Music is life who would love music it is a great thing for human being."

All schools fared well in OES question 8 with 86 % answering the "choose an answer" part correctly. Again this was not well reflected in the written answers or drawings, with only very occasional reasonable explanations being given, like the one below from a rural student:

> (R) "When 13 vocal cords change and become thicker voice comes out lower because vocal cords are thicker."

All schools also fared well in OES question 9 with 75 % answering the "choose an answer" part correctly. Once again though this was not well reflected in the written answers or drawings, with some confusion evident in the role of air in transmitting sound, like in the example below from a rural student:

> (R) "And a lot of wind in space cause sound not to be heard."

In general, written answers to OES questions did not provide good examples for comparing the original with the subsequent study or shedding light on new knowledge learnt from the ER (R-C), mostly because they were so poorly answered and even the reasonable answers were more descriptions than explanations. (Cooke and Breedin, 1994)

As revealed in "choose an answer" questions

These questions (indicating R-C ability) were mostly better answered in the subsequent study, with all schools improving their average mark (for the 5 questions): urban by 14 %, township by 7 % and rural by 10 %. Urban results were very similar (original to subsequent study) for Q 5 – 8 but 86% of the group answered Q 9 correctly, compared with just 40 % before. Township results were very similar to before, with 14 % fewer students answering Q 6 correctly than before, but 23 % more answering Q 8 correctly and an 11 % improvement for Q 9. Rural results were also similar on the whole, but with 29 % and 23 % improvements for Q 6 and 9 and a 23 % decrease for Q 7.

In summary, the only question to show a net decrease (original to subsequent) was Q 7 ("Why does a loud sound hurt your ears?") but this was just 4.8% across all three groups. No changes were made to the show which would assist in the answering of this question. Q 8 ("Why do men have deeper voices than women?") saw an improvement of 17%, perhaps reflecting the new slides and demos dealing with voice changes from boys to men. Q 9 was most improved with 27 % increase across all three groups. Two new demos (cellphone in lunch box and bell in vacuum) were added to the show dealing with this issue ("Can one hear sounds in outer-space") and this improvement reflects better student understanding of this question in the light of those demos – confirmed by the appearance of those demos in student drawings.

As revealed in quantitative probes

While the subsequent results for the three sections of the OES questions are interesting, they were not different enough from the original study to allow us to make clear comparisons – especially as to whether students were more able to deal with their difficulties related to sound and waves. An attempt to do that and to break it down for the different school groups follows.

Table 6.1 in section 6.1 simplified the difficulties table (original table was 2.2) to only those which were directly encountered in the student data from the original study, or could be expected to be changed by the content of the music show. Table 6.1 also indicated what intervention or modification would be instituted in the show to attempt to counter that difficulty.

An extensive analysis of the *quantitative* measures which probe these difficulties is presented in appendix p. Firstly the Multiple Choice and Open Ended survey questions directly probing each difficulty were identified, as well as the options for the correct and wrong answers, as in table 9.1 below. Then the percentage of students in each group choosing each option was presented and a ratio of correct/wrong calculated, for the pre-test and the post-test. (One would expect this to be < 1 for the pre-test and > 1 for the post-test). Finally a ratio of these ratios (post/pre) is compared for the subsequent and original studies to provide a crude measure of improvement between the studies.

Table 9.1: Quantitative questions probing each student difficulty

DIFF	ITEM	WRONG/ CORRECT	MC Q-NO.	OES Q-NO.	CHOICE
1	Sound travels as	TRANSVERSE	2		A
		LONGITUDINAL	2		C
6	Sound travels	THOUGH VACUUM		9	A
		NEEDS MEDIUM		9	B
8	Volume	RELATED TO FREQ.	6		D
		RELATED TO AMPL.	6		A
	Pitch	RELATED TO AMPL.	7		A
		RELATED TO FREQ.	7		D
8	Volume Unit	HERTZ	4		A
		DECIBEL	4		B
	Pitch Unit	DECIBEL	5		B
		HERTZ	5		A
9	Sound is fastest	THROUGH AIR	3		C
		THROUGH SOLIDS	3		D
		THROUGH AIR		5	B
		THROUGH SOLIDS		5	A
13	Music has	LOW VOLUME	8		A
		SMOOTH WAVE	8		C
14	Longer objects	FAST OR HIGHER	9		A, B, C
		SLOW AND LOWER	9		D
15	Sound box makes	MUSIC CLEARER		6	B
		MUSIC LOUDER		6	C
16	Vibration is	WAVE	1		A
		REGULAR MOVEMENT	1		C
17	Sound travels in	ELECTROMAGNETIC W	2		B
		LONGITUDINAL W	2		C

In order to quantify this comparison, a simple Likert scale was used to score:

VERY MUCH WORSE	MUCH WORSE	WORSE	IDENTICAL OR SAME	BETTER	MUCH BETTER	VERY MUCH BETTER
-3	-2	-1	0	+1	+2	+3

The results (detailed in appendix p) comparing improvements original to subsequent studies for all schools, rural, township and urban are shown in table 9.2 below

Table 9.2: Improvements in combatting student difficulties: original to subsequent test

DIFF	ITEM	QUEST-ION	ALL CHOOLS ORIG-SUBS	RURAL ORIG-SUBS	TOWNSHIP ORIG-SUBS	URBAN ORIG-SUBS
1	Sound travels as	MC 2	BETTER	BETTER	MUCH BETTER	WORSE
	TRANSVERSE		1	1	2	-1
6	Sound travels	OES 9	V. M. BETTER	V. M. BETTER	V. M. BETTER	V. M. BETTER
	THOUGH VACUUM		3	3	3	3
8	Volume	MC6	IDENTICAL	V. M. BETTER	V. M. WORSE	SAME
	RELATED TO FREQ.		0	3	-3	0
	Pitch	MC 7	WORSE	MUCH BETTER	MUCH WORSE	MUCH WORSE
	RELATED TO AMPL.		-1	2	-2	-2
8	Volume Unit	MC 4	MUCH WORSE	MUCH WORSE	MUCH WORSE	V. M. WORSE
	HERTZ		-2	-2	-2	-3
	Pitch Unit	MC 5	BETTER	MUCH BETTER	V. M. BETTER	SAME
	DECIBEL		1	2	3	0
9	Sound is fastest	MC 3	SAME	WORSE	V. M. BETTER	(SAME)
	THROUGH AIR		0	-1	3	0
	Sound is fastest	OES 5	SAME	SAME	SAME	BETTER
	THROUGH AIR		0	0	0	1
13	Music has	MC 8	(SAME)	(SAME)	WORSE	BETTER
	LOW VOLUME		0	0	-1	1
14	Longer objects	MC 9	IDENTICAL	MUCH BETTER	SAME	SAME
	FAST OR HIGHER		0	2	0	0
15	Sound box makes	OES 6	IDENTICAL	MUCH BETTER	WORSE	BETTER
	MUSIC CLEARER		0	2	-1	1
16	Vibration is	MC 1	MUCH BETTER	SAME	SAME	V. M. BETTER
	WAVE		2	0	0	3
17	Sound travels in	MC 2	(SAME)	MUCH BETTER	MUCH BETTER	MUCH WORSE
	ELECTROMAG. W		0	2	2	-2
		SCORE	**4**	**14**	**4**	**1**

From the table above, and in terms of addressing our difficulties with respect to sound and waves, the following conclusions may be drawn. When all schools are considered together (fourth column), there is little change from the original to the subsequent study. The difficulties where positive change was noted are:

- 1) an improvement of 1 point on our Likert scale. Here a new demonstration was introduced using a slinky sling and other animations to illustrate different types of waves
- 6) an improvement of 3 points on our Likert scale. Here a video was shown of a bell ringing silently in an evacuated bell jar.
- 8) a slight improvement in knowing the unit for pitch was shown, but the knowledge of the unit for volume dropped quite a lot
- 16) an improvement of 2 points on our Likert scale. Again the new wave type demonstrations mentioned above assisted.

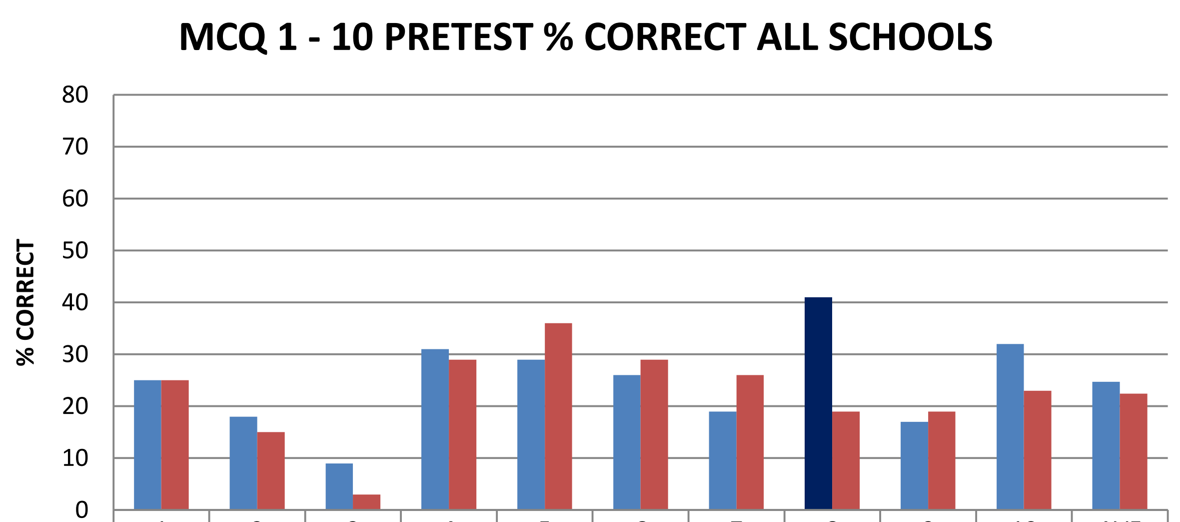
It is pleasing to see that the difficulties best addressed appear to be those where new demonstrations were added.

When considering individual school groups, the urban school showed very little change (1 point), the township school some change (4 points), but the rural school a marked change (14 points) with improvements in addressing almost every difficulty. According to our scale (comparing the subsequent to the original study), there was one instance of performance getting much worse, one getting worse, 3 staying the same, 6 getting better or much better and 2 very much better. Thus the new improved show seemed to assist the rural group most in terms of dealing with student difficulties. These are very crude measures however, and though they point to some improvement for the rural school, we need to consider what change occurred in more detail by looking at the MCQ individually. This section seems to indicate that the probes have improved but the show still needs work to show improvement for all three groups.

9.3 Prior knowledge before and learning during the show (RQ 2c)

The previous sections have mentioned that the interaction with students was too limited and shallow to show meaningful change in students' ability to reason with the ER (R-M) as reflected in answers to the OES. This section focusses on what students knew before the show(C) and whether there was any improvement after the show (R-C) in terms of conceptual knowledge. In particular, whether there was any discernible difference between the results from the subsequent study, compared with the original one. Let us consider the results of the 10 multiple choice questions (MCQ) which were offered both before and after the show. Detailed results of these tests are given in appendix j and only highlights will be given here.

When considering all three school groups together, the subsequent pre-test results showed similarly weak prior knowledge as the original one. Figure 9.14 below compares the percentage of correct answers for each question for the two studies; with the graph highlighted if the two studies' scores differed by more than 10%.

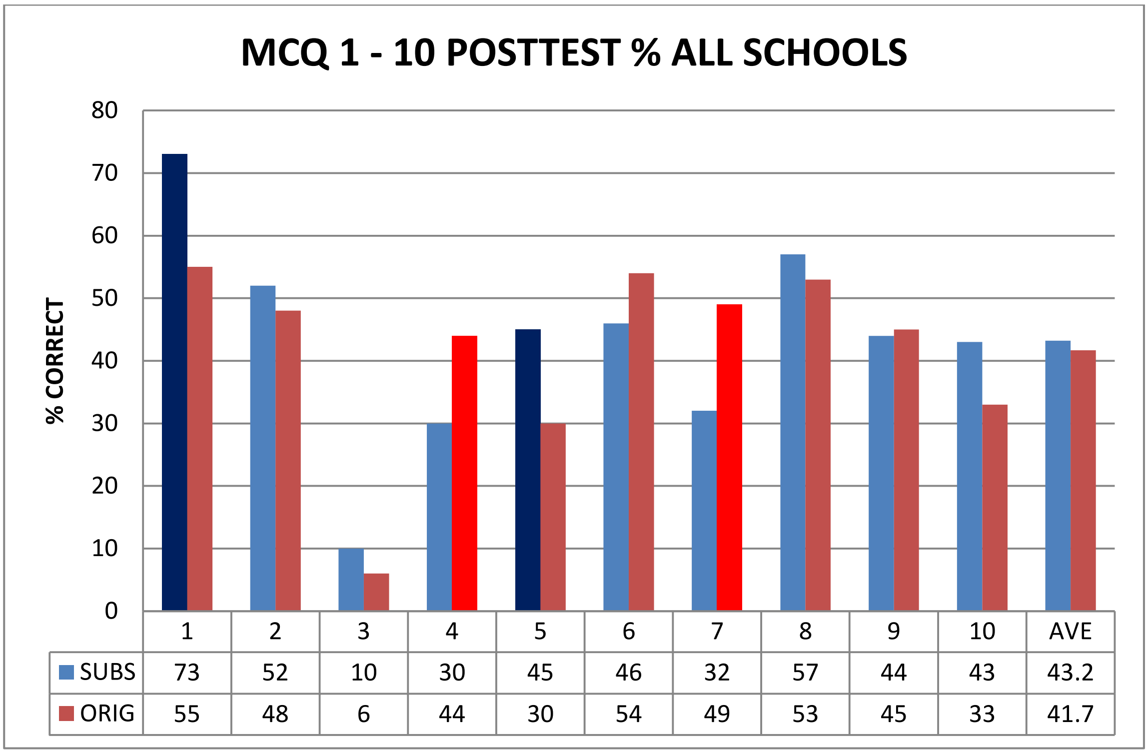


	1	2	3	4	5	6	7	8	9	10	AVE
SUBS	25	18	9	31	29	26	19	41	17	32	24.7
ORIG	25	15	3	29	36	29	26	19	19	23	22.4

Figure 9.14: Graph of percentage of correct student answers per MC question: pre-test comparing original to subsequent studies

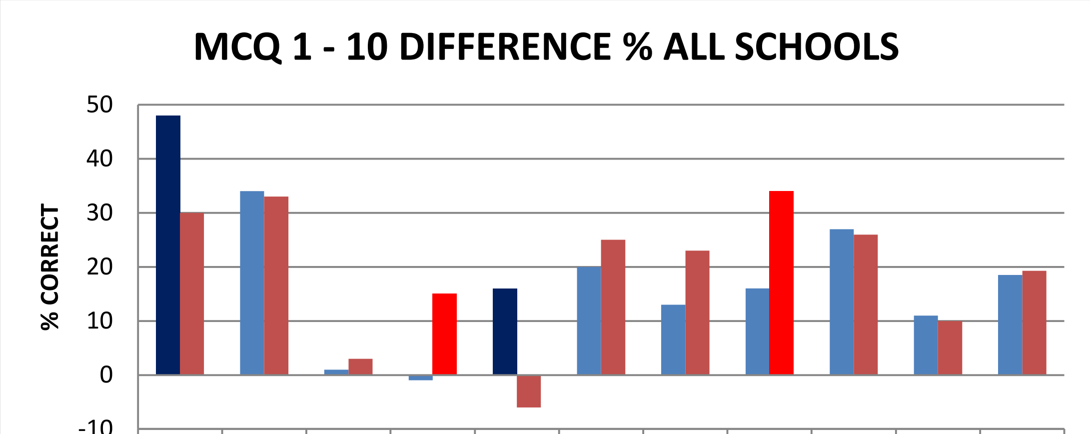
One can see that the two studies' scores differ very little and the average pre-test scores differ by only 2.3 %. The only real difference is in Q 8 where the subsequent score was 22 % higher than the original one. This question was rewritten to try to make it clearer, and this better score may be a result of that, or perhaps just a random effect (as Q 9, similarly rewritten, showed no improvement in the pre-test.) This similarity of scores and average suggests that we can regard the subsequent cohort of students as very similar to the original one in terms of prior knowledge of sound and waves.

The graphs below compare post-test percentages and differences (pre to post-test) for the original and the subsequent study, again with differences of more than 10 % highlighted. Question 1 (what is a wave?) showed improved knowledge gain in the subsequent study and quite a lot had been changed in the show to try to make this clearer. Q 4 and 5 (units for volume and frequency) showed gains and losses compared with the original study, and it is not clear just why this is the case. While Q 8 showed more *gains* for the original study, the better subsequent pre-test mark reported above meant that post-test marks were about the same. It can be seen from the averages that both post-test scores and differences show little change between these two studies considering all schools together. This is a little disappointing having improved the show, but we need to examine the individual results more carefully.



	1	2	3	4	5	6	7	8	9	10	AVE
SUBS	73	52	10	30	45	46	32	57	44	43	43.2
ORIG	55	48	6	44	30	54	49	53	45	33	41.7

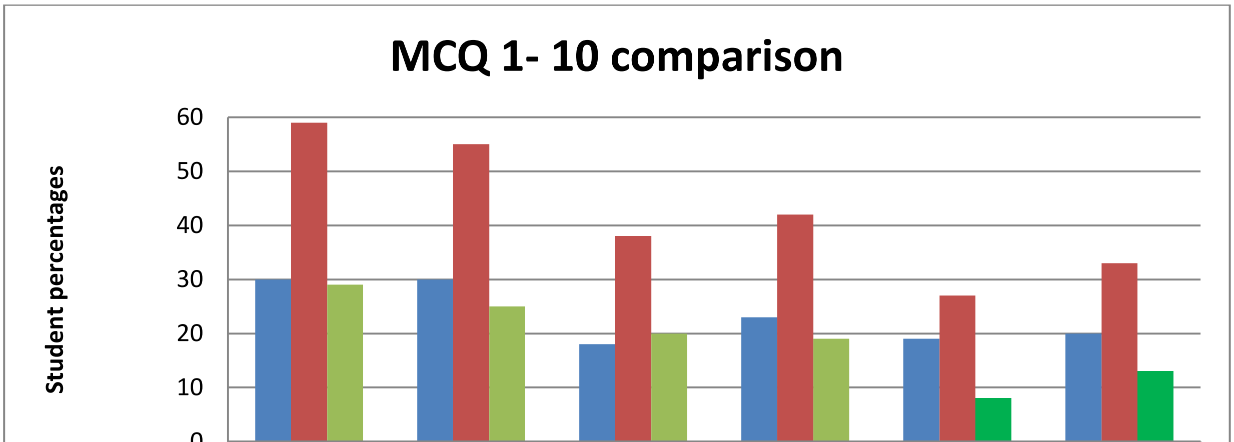
Figure 9.15: Graph of percentage of correct student answers per MC question: post-test comparing original to subsequent studies



	1	2	3	4	5	6	7	8	9	10	AVE
SUBS	48	34	1	-1	16	20	13	16	27	11	18.5
ORIG	30	33	3	15	-6	25	23	34	26	10	19.3

Figure 9.16: Graph of percentage difference per MC question: post-test minus pre-test comparing original to subsequent studies

When separated into individual school groups, the following pattern emerges: Figure 9.17 below shows pre-test, post-test and differences (as a percentage) for each school group, comparing the original with the subsequent study. For urban it can be seen that pre-test percentage was identical (original to subsequent), post-test and therefore difference was slightly down. With township, pre-test and post-test percentages were higher (5 % and 4 %), and the difference about the same (down 1 %). With the rural group, however the difference had increased from 8 % (original study) to 13 % - the only group to show this improvement. While the rural marks are still less than others, it appears that the improved show had a positive effect on their group, but not on the others.



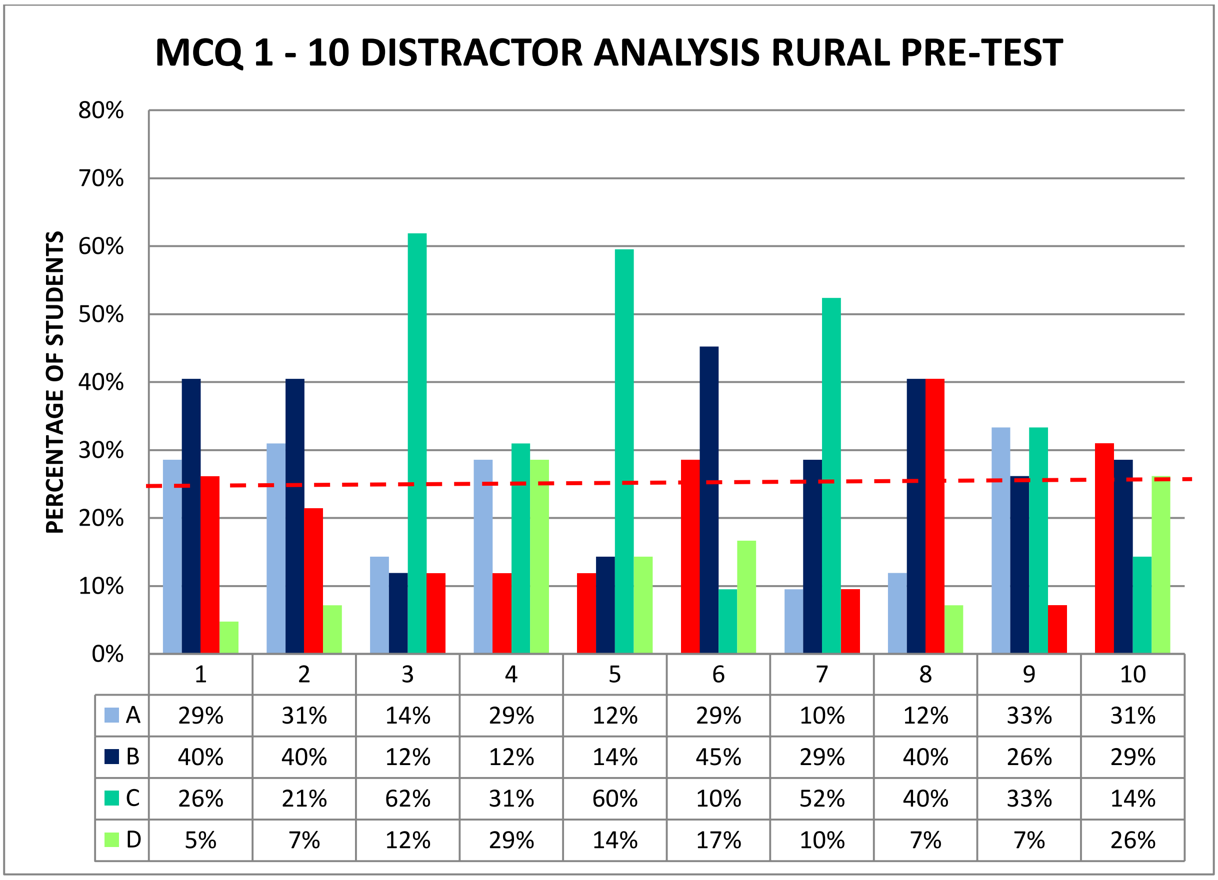
	ORIG URB	SUB URB	ORIG TOWN	SUB TOWN	ORIG RUR	SUB RUR
PRE-TEST %	30	30	18	23	19	20
POST-TEST %	59	55	38	42	27	33
DIFF %	29	25	20	19	8	13

Figure 9.17: Graph of MC test scores in school groups comparing original to subsequent studies

In addition to looking at absolute score per question, it is interesting to look at questions where the most chosen option was in fact the correct one and compare the original study with the subsequent one. Table 9.3 below gives the number of questions (out of 10) where the most chosen answer was the correct one. It can be clearly seen that urban and township groups only really show improvements in the pre-test, but rural shows a similar pre-test but greatly improved post-test. While there are some rural children who still are unable to learn from the show, those who can learn appear to be doing so better in the subsequent study. This is further shown in figures 9.18 and 9.19.

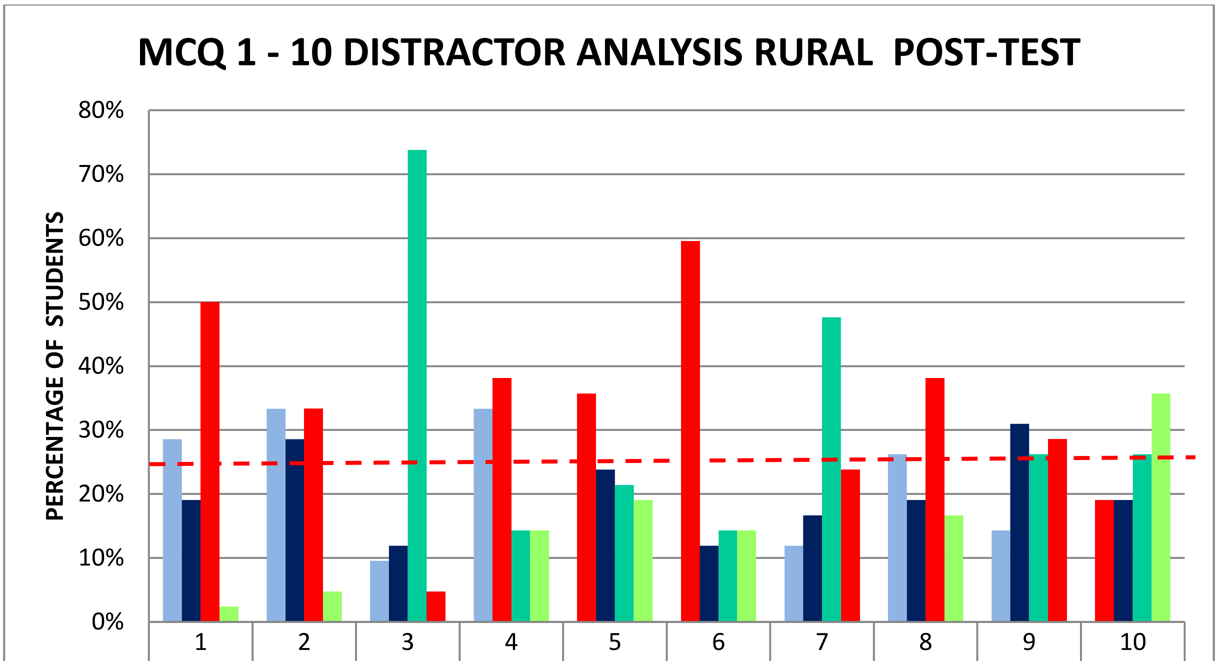
Table 9.3: Number of MC questions where most chosen option was correct

	ORIGINAL STUDY		SUBSEQUENT STUDY	
GROUP	PRE-TEST	POST TEST	PRE-TEST	POST-TEST
URBAN	4	8	5	7
TOWNSHIP	0	6	2	6
RURAL	2	2	2	**6**
ALL SCHOOLS	2	9	2	7



	1	2	3	4	5	6	7	8	9	10
A	29%	31%	14%	29%	12%	29%	10%	12%	33%	31%
B	40%	40%	12%	12%	14%	45%	29%	40%	26%	29%
C	26%	21%	62%	31%	60%	10%	52%	40%	33%	14%
D	5%	7%	12%	29%	14%	17%	10%	7%	7%	26%

Figure 9.18: Distractor Analysis: Rural pre-test, subsequent study (Key is coloured red)



	1	2	3	4	5	6	7	8	9	10
A	29%	33%	10%	33%	36%	60%	12%	26%	14%	19%
B	19%	29%	12%	38%	24%	12%	17%	19%	31%	19%
C	50%	33%	74%	14%	21%	14%	48%	38%	26%	26%
D	2%	5%	5%	14%	19%	14%	24%	17%	29%	36%

Figure 9.19: Distractor Analysis: Rural post-test, subsequent study (Key is coloured red)

Considering some individual questions: the incorrect conception held in MCQ 1 (b), that a wave is a movement in a straight line, was corrected after the show with a good increase in choices of the correct answer (c). MCQ 2 had seen no change pre- to post-test in the original study, but now the incorrect choice (b) describing sound waves as electromagnetic waves was less after the show, and the correct choice (c) somewhat improved. MCQ 4, 5 and 6 saw no improvement in the original study (with 6 getting worse) but now there was great improvement in all three. Confusions around volume and pitch (d 8) were corrected in all these questions, but not in MCQ 7, where the new distracter (c) that pitch is related to smoothness of the wave-shape was most popular in both tests. While MCQ9 was not comprehensively corrected, it was pleasing that the two “higher” choices were selected in the pre-test (a and c) and the two “lower” choices (b and d) in the post-test. It may be suggested that the rewriting of that question made it more understandable to the rural students. The concept of ultrasound in MCQ 10 is still not well understood by this group.

When these graphs above are compared, though, it can be seen that real progress has been made from pre- to post-test in most questions – especially when compared to the same graphs from the original study (figures 5.33 and 5.34) where almost no progress was made. Finally the increased confidence to attempt drawings by the rural group reported in section 7.4.2 further confirms the improvement in this group.

While the overall mark for the urban post-test dropped slightly, there were improvements in MCQs 1 and 10 compared with the original study. Section 5.3 reported that urban students incorrectly chose option (a) “a vibration is a wave” in both the pre- and post-tests in the original study. It was again chosen by 57% of students in the subsequent pre-test, but only 14% in the post-test, while the correct answer (c) increased from 19% to 86% of students. Apart from the comparison question (MCQ 3) this was the only question in the original study where the urban school did not show improvement (pre- to post-test) so it pleasing to see that this has been sorted out –especially when one considers the number of measures introduced into the show to clarify students’ ideas of what a wave is (table 6.1). MCQ 10 was thoroughly discussed before (section 5.3) showing that urban students chose the two frequency options (a and b) in the original test but were not sure which was correct. In the subsequent test the correct option (a) was clearly the first choice. What was of concern were that MCQs 4 and 5 were both answered mostly correctly in the pre-test, but then mostly incorrectly in the post-test. This happened with MCQ 4 for the township group as well but not for the rural group. It appears that some aspect of the modified show confused the urban group as to the units of volume and frequency. It is not yet clear what the cause of that confusion was.

For the township group, the improvement pre- to post-test was about the same as before. As mentioned above MCQ 4 got worse pre- to post-test, but MCQ 5 showed good improvement where there had been none in the original test. MCQ 6 and 7 showed no improvement where there had been good improvement in the original test. Again it is not clear what in the show caused this confusion which was not there for the rural group.

9.4 Discussion of results (RQ 3 and 4)

In summary for this section, the subsequent study showed very few clear gains for the urban and township groups, as revealed in both the MCQ pre- and post-tests (section 9.3, C and R-C), as well as the OES (drawings, written answers and "choose an answer" (see section 9.2, R-M). It was pleasing to see the one question MCQ 1 clearly improved for the urban group but puzzling to see others (MCQ 4 and 5) get worse. While the probes appear to work better (RQ 4) the show still needs some modifications. The rural group, who did not fare well in the original study, showed measurable improvement compared with the previous study, especially in pre- to post-test gains in the MCQ (R-C, where urban and township groups did not improve). Comparable responses to attitudinal questions (MCQ 11 – 13 see section 9.1.1) indicate that student attitude (RQ 2a) was not a major factor in making the subsequent group perform differently from the original one. In addition, very similar pre-test scores for the MCQ show that the original and subsequent groups had similar prior knowledge (C) of the subject matter. It would appear that the redesigned show (RQ 3) and probes (RQ 4) assisted the rural group in terms of gains in conceptual knowledge, but not really the other two groups. (This is commented on further in Chapter 11). This is meaningful though, in that the original show and probes worked well for urban and township students, but not for rural students. We now have (with the subsequent show and probes) instruments which we can confidently use across all three groups.

Chapter 10: Final study: isiZulu show

10.1 Final redesign of the show: isiZulu show (Stage 4: 2nd iteration)

After receiving some encouraging results showing improvements for the rural group based on the modified show (and testing instrument), one final change was attempted to see if further improvement was possible for this group. In the light of:

- language concerns raised in sections 2.4; 3.1.2; 3.3; 3.4.1; 5.4;
- the language issues affecting validation in section 4; and
- the written answers and results of the English Language Test (appendix d):

I decided to test what effect the language in which the show was presented would have on student results. After much discussion and debate with (isiZulu-speaking) colleagues and my supervisors, I decided to have the *show* translated into isiZulu (affecting the nature of the ER – M), but keep the *probes* in English, for the following reasons:

1) As I wish this PhD to be a practical study of actual practice in Science Centres which I trust will culminate in the generation of a useful tool for SC staff to use (in addition to all the other results of course), and as schooling ends in a matric exam which is written in English, I believe that the study which will best fit the current schooling situation and be most useful is thus one with instruction (the show) in isiZulu and the testing (the survey) and student responses, in English.

2) As much of the survey tests the extent to which students have come to terms with technical language (like frequency, amplitude, longitudinal waves etc.) for which no direct translations exist in isiZulu, it would be difficult to conduct the survey in isiZulu. (Many of the isiZulu versions of words like these are actually explanatory phrases, which negates a question asking for them to be explained)

3) Student written responses to questions need to be coded by myself and checked and recoded to test for inter-coder reliability by my three supervisors, none of whom speaks isiZulu. It would be impossible to do this reliably with a translator, especially with distant supervisors.

4) Changing the language of both the intervention (the show) and the instrument (the survey) at the same time will make it very difficult to know which was responsible for any gains (or conceivably losses) we might encounter.

5) In order to compare the results of an isiZulu show with previous results (from PhD June testing and MSc Nov 2011 testing), the testing needs to be done in the same way.

6) In future I hope to extend the study internationally, getting overseas colleagues to use the identical instrument with students at their centres. Again this would not be possible with a non-English instrument.

7) Zuma and Dempster (2008) showed that performance amongst isiZulu-speaking children was not significantly improved if the probe items were translated into isiZulu (even though students revealed a very positive attitude towards tests in isiZulu).

Issues of the language of instruction and testing are obviously extremely contentious and merit much more attention than this thesis will allow.

The translation of the slides was done by one of our student interns (mentioned in section 7.1) and was checked by three other isiZulu speaking science teachers. The process was not straightforward and served to underline the practical difficulties involved in translating science in South Africa. One example of these difficulties is in the translation of words I wanted to keep distinct to avoid confusion between pitch and volume (d 8):

- High and low for pitch
- Loud and soft for volume

The first translation yielded the same words ("phezulu" and "phanzi") for both "high and low" and "loud and soft", which could easily lead to confusion. Much further investigation and discussion led to the following formulation in table 10.1 below:

Table 10.1: Examples of isiZulu translations

English Word	isiZulu Equivalent
High sound (pitch)	Umsindo Ohlabayo
Low sound (pitch)	Umsindo Ondondayo
Loud sound (volume)	Umsindo Omemezayo
Soft sound (volume)	Umsindo Ohlebezayo

Many other similar translation challenges had to be overcome. While presenting in isiZulu the English technical terms (like "frequency") were carefully introduced. An example of an isiZulu slide from the show is in figure 10.1 below.

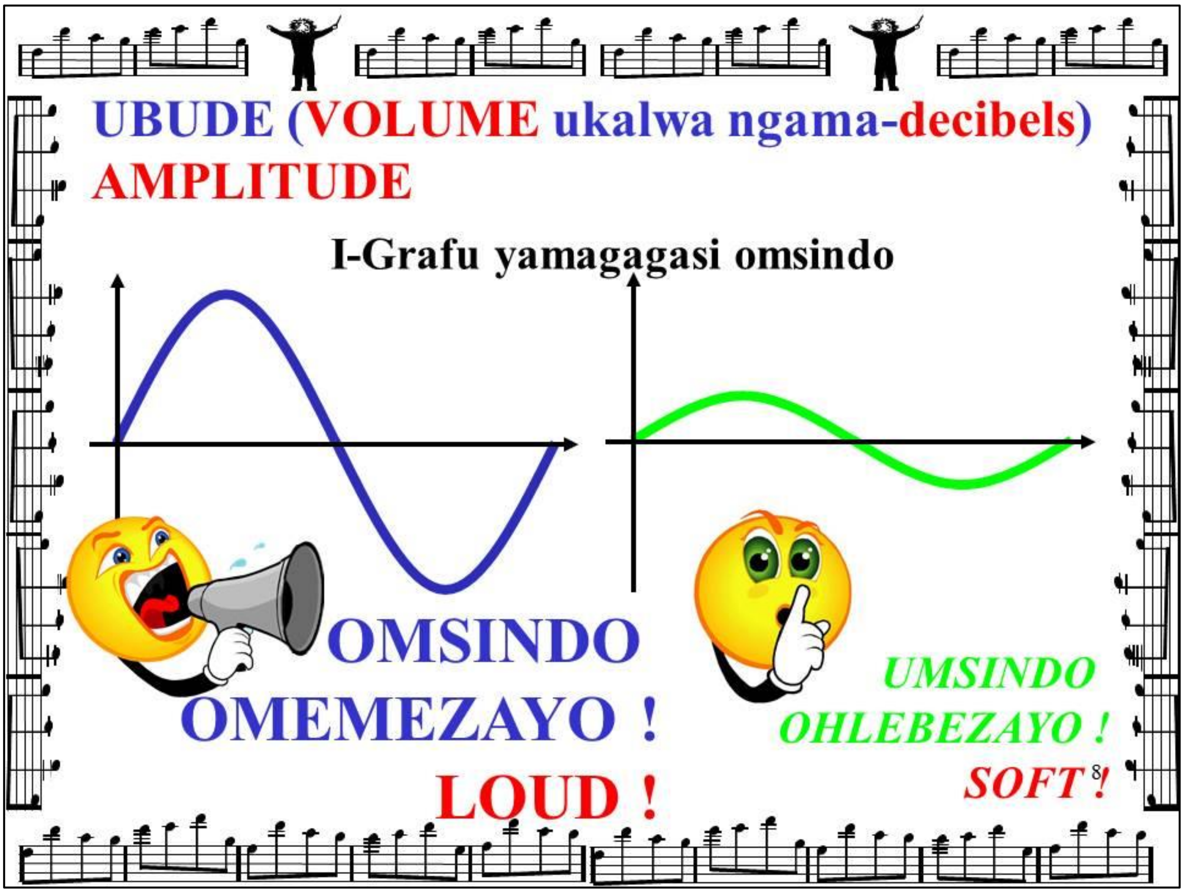


Figure 10.1: Example of a PowerPoint slide from the isiZulu show.

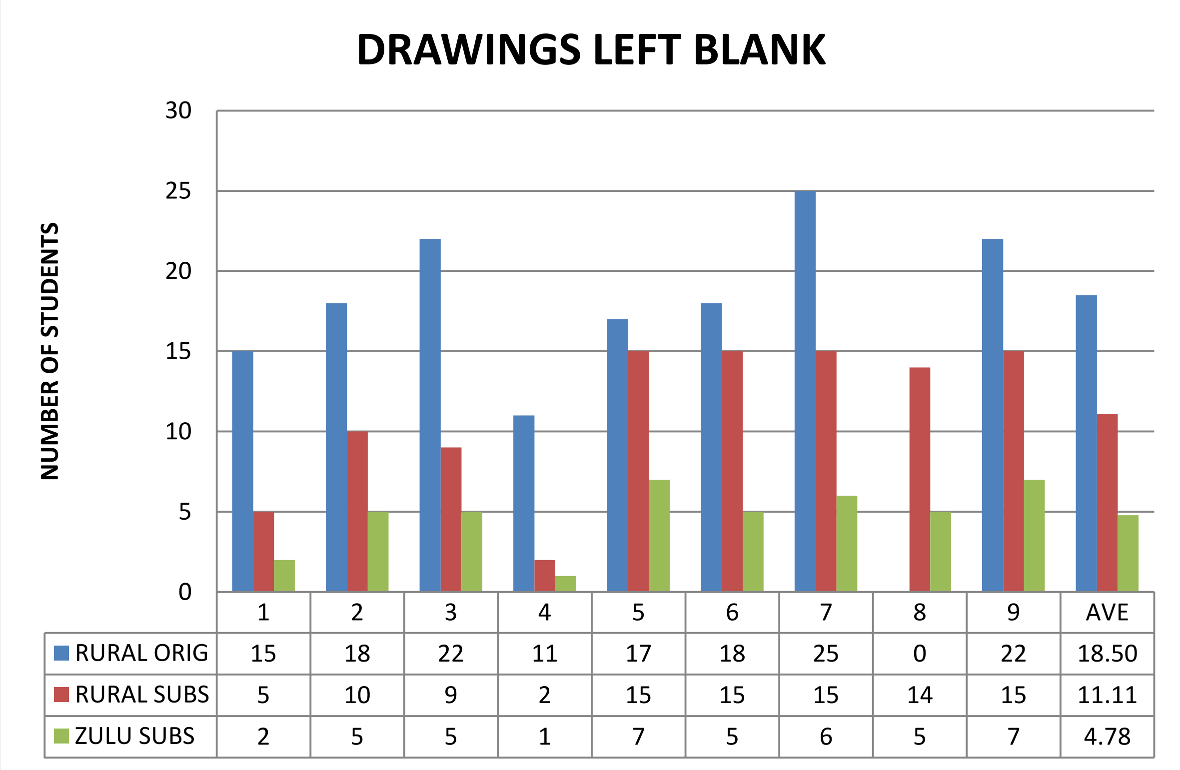
Apart from changes to the words on the slides, the whole show was presented in isiZulu, but the pre- and post-tests presented and written in English (although instructions for these were given in isiZulu). As the probes remained unchanged, no further validation of these is presented for this group but some issues will be commented on in the results section below (section 10.3).

10.2 Details of the retest: final show (Stage 3: 3rd iteration)

As mentioned above an isiZulu show was developed to attempt to test what the effect of language of presentation would be on the results in the same English medium probes. This testing was done about 4 months after the subsequent tests, with a group of 42 different Grade 9 pupils from the same rural school as used previously. Conditions and day-programme were identical, except that the show was presented in isiZulu. Separate ethical clearance was obtained for this modification (appendix e).

10.3 isiZulu group results (Stage 3: 3rd iteration)

OES results in the three sections for the isiZulu group were not very different than for the subsequent rural group, except that an enhanced confidence to attempt drawings was shown in the vastly reduced number of drawings left blank. Figure 10.2 below compares number of drawings left blank for the original rural group, subsequent rural group and isiZulu group:



	1	2	3	4	5	6	7	8	9	AVE
RURAL ORIG	15	18	22	11	17	18	25	0	22	18.50
RURAL SUBS	5	10	9	2	15	15	15	14	15	11.11
ZULU SUBS	2	5	5	1	7	5	6	5	7	4.78

Figure 10.2: Drawings left blank by the three rural groups

The drawings themselves were very similar to those of the subsequent rural groups, and the written answers again were almost unusable because of the very weak English language ability of the students. As reported before, it appears the scope of this ER is too limited to affect ability to reason visually with the ER to a great extent.

The difference in results for this group was most marked in the MCQ post-test (R-C). We have already seen in fig 9.17 how the rural group improved from the original to the subsequent study. Figure 10.3 below compares pre-test and post-test percentages, as well as differences (improvement post minus pre) for all four groups in the subsequent study.

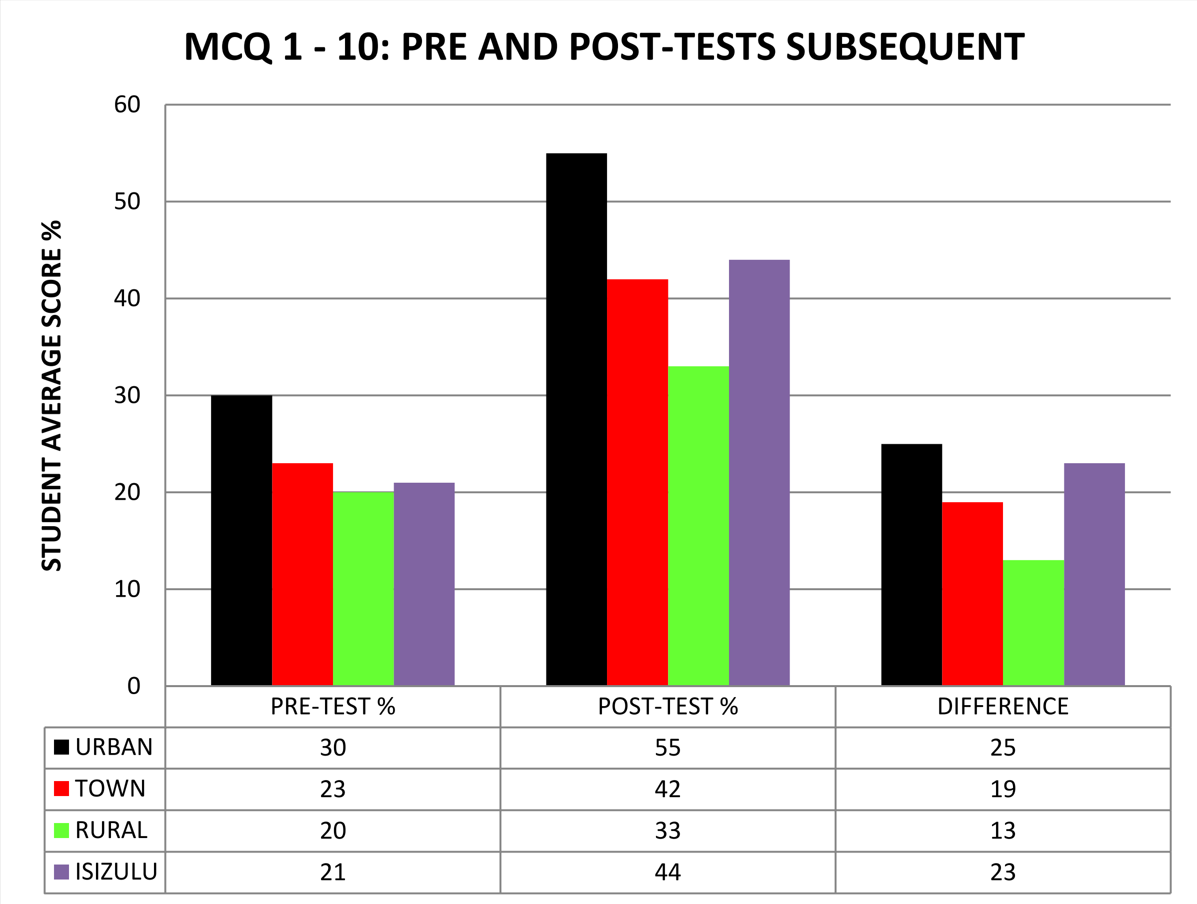


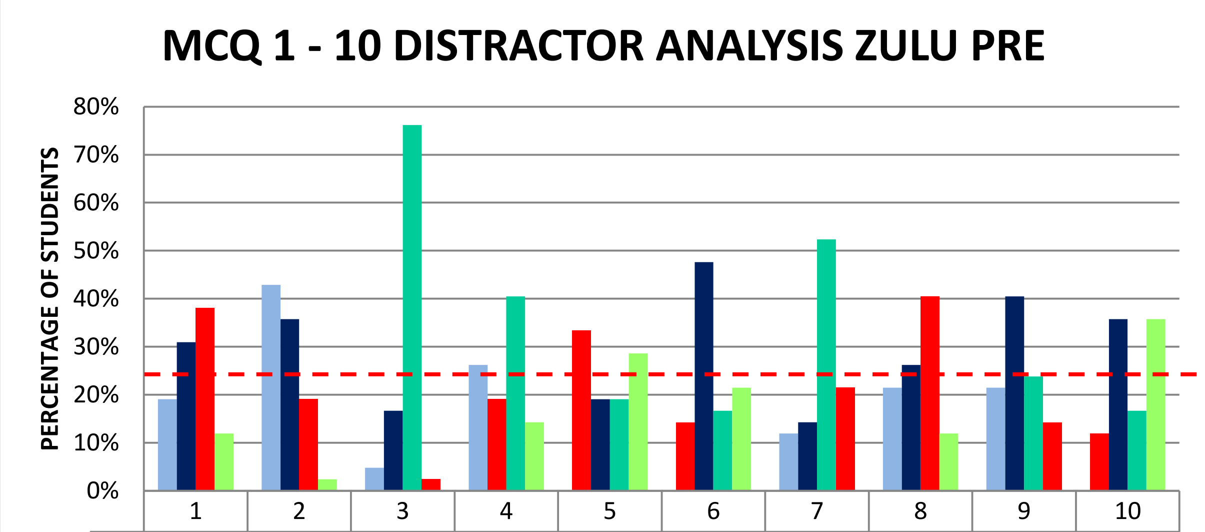
Figure 10.3: Graphs of MCQ's group averages for four subsequent groups

It can be clearly seen what a difference home language instruction made to the isiZulu group. Their pre-test mark (C) is almost identical to that of the rural group (21 and 20 %), but their post-test mark (R-C) (44 %) and difference (23 %) is much better, and even better than the township group. Their t-test value comparing pre- to post-test scores was even lower than before confirming that this improvement was not by chance. The standard deviations of pre- and post-test scores were very similar to the rural ones indicating an increase in mean score, rather than just a greater spread. Average discrimination index for all 10 questions was also improved.

Table 10.2 Statistics for rural and isiZulu pre- and post-tests (MCQ)

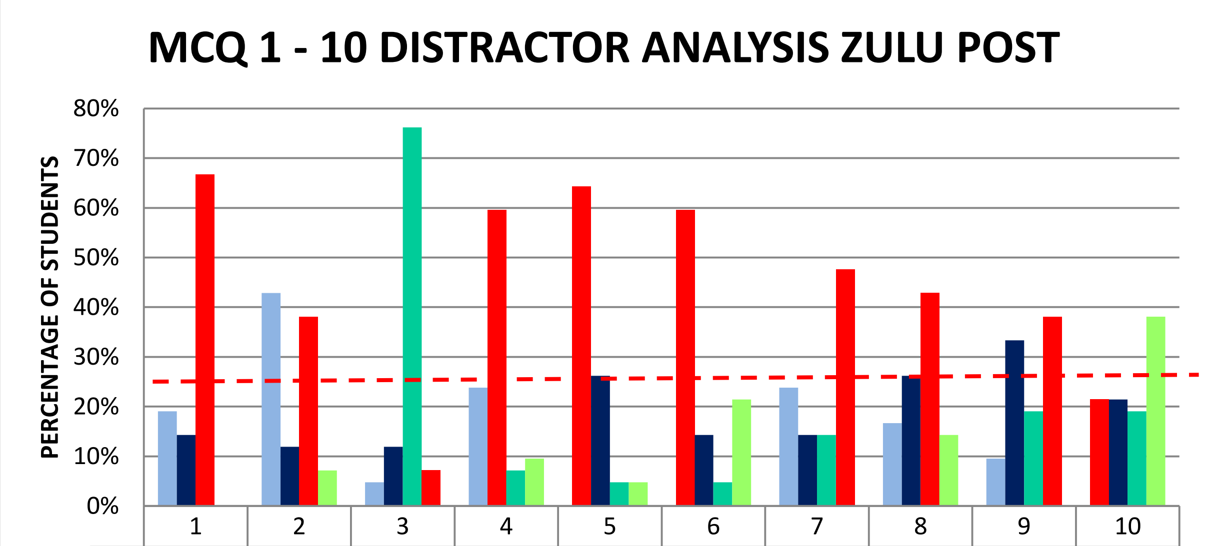
	AVERAGE %	STANDARD DEVIATION %	T-TEST SCORE	AVERAGE DISCRIM. INDEX
RURAL PRE-TEST	20.0	13.8	2.4 E-5	0.51
RURAL POST-TEST	33.1	17.3		
ISIZULU PRE-TEST	21.4	13.2	2.1 E-9	0.57
ISIZULU POST-TEST	44.5	18.5		

Figures 10.4 and 10.5 below show the pre- and post-test distracter analyses for the isiZulu group, highlighting the improvement in *every* question and in starting with only 3 questions where the most common answer was correct and ending with 7.



	1	2	3	4	5	6	7	8	9	10
A	19%	43%	5%	26%	33%	14%	12%	21%	21%	12%
B	31%	36%	17%	19%	19%	48%	14%	26%	40%	36%
C	38%	19%	76%	40%	19%	17%	52%	40%	24%	17%
D	12%	2%	2%	14%	29%	21%	21%	12%	14%	36%

Figure 10.4: Distractor Analysis: isiZulu group pre-test (Key is coloured red)



	1	2	3	4	5	6	7	8	9	10
A	19%	43%	5%	24%	64%	60%	24%	17%	10%	21%
B	14%	12%	12%	60%	26%	14%	14%	26%	33%	21%
C	67%	38%	76%	7%	5%	5%	14%	43%	19%	19%
D	0%	7%	7%	10%	5%	21%	48%	14%	38%	38%

Figure 10.5: Distractor Analysis: isiZulu group post-test (Key is coloured red)

The pre-test distracter analysis shows a number of questions (aside from MCQ 3) where students held incorrect prior conceptions. In MCQ 2, option (a) that sound travels in a transverse wave (d 1) was most popular and was not corrected in the post-test after the show, nor in their drawings for OES Q 1 – 3. MCQ 6 (b) was popular in the pre-test, that volume is related to wave-speed, but was much improved after the show. Similarly, the new distracter, 7 (c) that pitch is related to smoothness of the wave-shape, was chosen in the pre-test but corrected in the post-test. This group showed little of the volume and pitch confusion of previous groups, as the questions (MCQ 4 – 7) dealing with these quantities all saw good improvement in the post-test. It is pleasing to note that the misconceptions 4 (a) and 7 (c) which were evident in figures 8.3 and 8.4 for the subsequent test have been addressed for most of this group of students. This seems to indicate that home language instruction was effective even for communicating the English terminology effectively. MCQ 9's option of (b) "faster and lower" was corrected to (d) "slower and lower" in the post-test. MCQ 10 saw the isiZulu group choose the two "lower" options (b and d) and this was not corrected in the post-test, indicating a difficulty understanding frequency.

What was of concern in the original rural group was that almost a quarter of students actually got *worse* pre- to post-test, indicating that they had little grasp of the questions. While this improved slightly in the subsequent rural group, figure 10.6 below shows that it improved even further for the isiZulu group, giving evidence that the show is now more understandable in answering RQ 3.

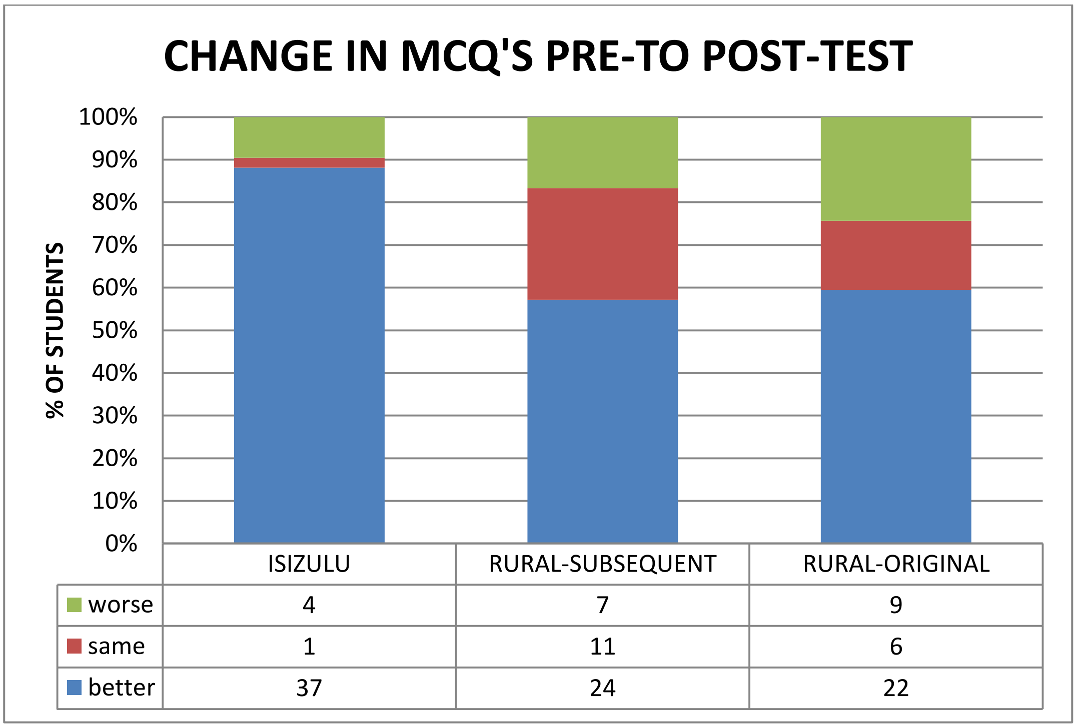


Figure 10.6: Graph of change in MCQ scores pre- to post-test for the three rural groups

Finally, to ensure that the better results for the isiZulu group were not merely the result of better English language ability, the same English language ability test (appendix d) applied to the *original* rural group was administered to this group during a school visit sometime after their visit to the Science Centre. The group available at the school numbered 30 (of the 42), and the following results (all marks in %) were obtained (compared with those from the *original* study rural group).

Table 10.3: English language test results comparison

	n	AVERAGE %	LOW %	HIGH %	STDEV %
RURAL-ORIGINAL	22	54.5	28	87	13.5
ISIZULU	30	48.5	23	81	13.9

The isiZulu group actually performed slightly worse in the English test than the original rural group so we can exclude better English language ability as the reason for their better performance in the MCQ's.

10.4 Discussion of results (Stage 3: 3rd iteration)

The extra test done with the presentation in isiZulu revealed further improvements (section 10.3) as shown especially in the MCQ where this group actually outperformed the township group and more than doubled their class average pre- to post-test, despite the test being in English, and their general English language ability being poor. Furthermore they showed that incorrect prior conceptions were effectively dealt with, including the understanding of English terminology like volume and pitch. This was less evident in their OES results, where they were still hampered in crafting written explanations by their very weak English language ability, but their confidence to attempt the drawings was much better.

Clearly the nature of the ER (M), in this case the language of instruction, was probably the most important factor in affecting the students' general ability to successfully interpret, visualize and learn from the ER (C-R-M) and an o=important contribution to answering RQ 3.

Chapter 11: Final Discussion and Conclusions

11.1 Major achievements of this study

The research data presented in this thesis permitted me to make the following inferences:

- That student attitude is affected only mildly by the type of school attended (5.1) and that all groups appeared to enjoy the science show equally (from MCQ 11 in the post-test, and fig. 5.2)
- That student conceptual and visual difficulties with respect to sound vary between school types, with urban schools tending to embrace a wave-based perspective of sound, rural a more object-based one and township schools more of a blended model (table 5.3)(Hrepic, 2010)
- That these difficulties (table 5.1) to some extent mirror those found in the literature (like difficulties d 1, 5, 6, 8, 9 and 13 – 16) but also include novel difficulties not seen before (d 17 – 19, observed from answers to MCQ 2 and OES 2 and 3) and that some of the difficulties recorded in the literature were not found from my data (d 2 – 4, 7, 10 – 12)
- That the major difficulty encountered by Grade 9 students with regard to sound, is the conception that sound travels with transverse waves (d 1 on table 5.1, revealed especially in drawings in OES 1 - 3)
- That, in line with conceptual change theory, conceptual and visual difficulties (for example the difficulty d1, mentioned above, that sound travels in transverse waves) are more deeply entrenched and are not easily changed with one short science show.
- That students attending a science show can achieve substantial learning gains, doubling their scores from pre-test to post-test (table 5.3) and with almost all incorrect prior conceptions remediated for most students (as in the discussions on MCQ 4 – 7 where difficulties d 8 on the confusion between volume and pitch were corrected for the majority of students in the post-test (fig. 4.4))
- That the learning gains do not apply when an item is deliberately excluded from the show (as for MCQ 3, fig. 4.4, reconfirmed in figs. 8.4 and 10.5), convincing us that learning can be linked to the show.
- That student prior knowledge and learning during a science show is dependent on the type of school attended, with urban schools performing better than township schools that in turn outperform rural schools (table 5.3). This mirrors the findings of the TIMSS studies (www.timss.bc.edu) linking achievement in science with the size of the nearest town.

- That the multiple choice probes developed in this study are effective in revealing evidence of students' thinking about sound across all groups but written answers are strongly dependent on English language ability (table 4.4) and are effectively unusable for the rural group under study. Drawings provide rich data if one can overcome student reluctance to attempt them.
- That modifications to the show can increase student learning but will tend to affect the weakest groups the most (table 9.2), as revealed by the rural group's overall increase in pre- to post-test in the MCQ (table 9.3). In addition new demonstrations on the nature of waves (e.g. slinky spring) led to great improvement on corresponding probes MCQ 1 and 2 (figure 9.19) which had not been seen in this group before the show was modified.
- That modifications to the probes (table 7.1) improved our ability to detect incorrect conceptions and to view the performance of the rural group - the weakest group - effectively. The effects of such modifications are more easily seen in the results from the quantitative probes (fig. 9.19) and in a higher discrimination index for this subsequent rural group (fig. 8.2).
- That the greatest improvement for the weakest group in terms of overall score increase (pre- to post-test MCQ), and confidence in attempting drawings, is achieved by presenting the show in their home-language (section 10.3). More specifically, MCQs 4 – 7 showed that this group (instructed in isiZulu) made great progress in sorting out the confusion between volume and pitch evident in the pre-test (Figs. 10.3 and 10.4).

11.2 Extent to which research questions were addressed:

RQ 1) What types of probes would effectively reveal evidence of student learning about sound and any related conceptual and visual difficulties after experiencing a science centre show *on the physics of sound*?

Science Centres are still relatively new in South Africa, and there have been few research studies done locally on them. The only evaluations happening in the centres are by and large "visitor surveys" of how the visit was enjoyed and experienced. There is, therefore, a great need to provide effective assessment tools for those working in science centres. This research question was comprehensively answered with a thorough study of many different probes which can be used to assess a science show. The probes were thoroughly tested and revised through two pilot studies (3.4.1 and 3.4.2) and the individual questions were scrutinized and critiqued by a panel of four experts (two in Physics and two in Education and assessment), as well as by peers at various physics conferences. The final survey went through an exhaustive validation process (chapter 4) before it was used to draw conclusions about the students (chapter 5). Student data was used for item analyses and statistical analyses of quantitative data, as well as triangulation and other validations of the qualitative data.

Validation (Chapter 4) was very successful with all of the original probes proving usable, though it was noted that they were not valid across all school groups. The instruments showed great validity for the urban and township groups but less valid for the rural group. The study proceeded to investigate what type of instrument (or modification of this instrument) would be most effective to assess this rural group, in answering RQ 4. Statistical results showed extremely high significance and an acceptable spread of results although again this varied between the different groups.

The various probes listed in table 3.2 in section 3.3 were initially used with 117 students from 3 very different school groups to produce the results of chapter 5. The list below summarises the fact that:

1. Multiple choice questions (appendix b) provided a quick and useful test of prior knowledge (C) and learning during the shows (R-C) by comparing post-test scores with pre-test. They appeared to work well across all school groups, probably because they only required reading, and not writing skills. Statistical indicators showed that they worked best for urban students, but still satisfactorily with the other groups. They provided quick evidence of measurable learning during the show (Abdalla, 2013).
2. Multiple choice attitudinal questions (appendix b) provided swift feedback across all school groups
3. Written answers (appendix c) provided rich feedback in terms of assessing student conceptions after the show (R-C), but only for urban students. Township, and especially rural, students were so hampered by poor writing skills that the questions did not work well for them at all. More research is needed on how best to assess these groups and on the generation of explanations (Cooke and Breedin, 1994).
4. Drawings (appendix c) provided fascinating insights into students' visualizations of sound after the show (R-M), but were limited by township and especially rural students' reluctance to attempt them. Students clearly lacked experience and skill in producing drawings, often providing *illustrations* rather than *explanations*. (Kose, 2008)
5. "Source of knowledge" questions provided average feedback and had some value in tying learning to the show, but were again hampered by township and especially rural students leaving them blank.

6. The English language test (appendix d) provided quick and useful feedback on students' general language abilities and worked well across all school groups, correlating well with the standard of written answers.

Again this feedback provides very useful advice to South African practitioners as all the other studies reported in the literature were performed outside the South African context. This initial feedback will be further qualified by comments addressing RQ 4 below, on how the probes were fine-tuned after the first iteration.

RQ 2) How do the students from the *urban, township and rural* school groups compare, after experiencing *the sound show*, in terms of:

a) General attitude;

Section 5.1 summarises the attitudinal responses from the students. All groups reported that they had found the show enjoyable which enabled me to eliminate that as a factor affecting relative performance. What was surprising was that the urban students, with the best resources at school, found school science to be least enjoyable, while rural students found it most enjoyable. Rural students reported that they had found the show more difficult to understand than the other groups, which was reflected in their weak scores in the post-test. Falk and Gillespie (2009) stress the role that emotion plays in free-choice learning and affirm that research in this area is extremely limited. It is clear that there is much more that can be probed in this regard which provides fertile soil for future studies – especially with the added value of introducing local data into the research. Packer and Ballantyne (2002) similarly confirm the important role motivation plays in free-choice learning at museums and science centres.

b) Conceptual and visual difficulties with respect to sound

The tables at the end of section 5.2 summarise my findings on student difficulties understanding sound, based on a synthesis of data from almost all of my probes. I found many of the conceptions from the literature to be present in the answers of the students to differing degrees, but also added a further three novel difficulties (17 – 19) at level 1 as I found these to occur regularly in the answers across all three groups (although not in the literature). In addition to showing which difficulties were present and which not, the table also shows which school groups especially showed that difficulty. Urban and township schools were fairly similar in terms of their difficulties, but rural somewhat different. This may be attributed to the fact that rural students were less likely to describe phenomena in terms of waves, preferring to use object-based descriptions; so they exhibited fewer of the wave-based difficulties, and more of the concrete object-based ones.

Our modified table Fig. 5.1 presented in section 5.2 provides an invaluable resource for local practitioners. It not only adds local data to the foreign studies, but even highlights which difficulties are more likely to be prevalent in which type of school group. The information in the revised table fills a much needed gap in the literature with locally relevant data and conclusions. It would now be even more valuable for comparative studies to be performed locally so that the table can be further augmented and improved. It is not yet clear whether these difficulties would be stable across varying contexts (McCloskey 1983) or whether we need to apply more of a "knowledge in pieces" approach (di Sessa, 2002) for the best understanding of these difficulties. Perhaps a further study which varies the contexts (Cooke and Reedin, 1994) may be appropriate.

Visual difficulties both prior (C) and in response to the ER (R-M) abounded across all student groups as revealed mostly in the drawings. It is doubtful to what extent these can be addressed in one 45 minute show despite attempts to improve the show (C-M) and so they are better tackled in a classroom situation. Conceptual change is difficult and this study indicates that more than a brief interaction with students is needed. In section 2.2 I listed Posner's (1982) conditions for students to take on the accommodation required for conceptual change:

1. Dissatisfaction with their existing conceptions and a realization of their inconsistencies;
2. A new conception which is intelligible and makes sense;
3. A new conception which appears initially plausible and makes more sense than the old one; and,
4. A new conception which suggests the possibility of fruitful new areas of inquiry.

In the light of this study I feel that we may be able to create the conditions for 3 and 4, but that 1 and 2 would need more contact time with students. Many of the students' visual difficulties with sound were related to their confusing representations (like graphs) of reality, and these were still evident after the show. As mentioned in section 2.2, Duit (2003) has provided many examples of conceptual change failing to occur and, therefore, we cannot assume it will be easily or quickly achieved.

c) Prior knowledge and learning relating to the show?

The groups differed in terms of prior knowledge (C), with the multiple choice pre-test showing the urban group clearly ahead of the other two. The multiple choice post-test showed clear evidence of learning during the show (R-C) for urban and township groups with each doubling their scores from pre-test to post-test, and with T-Test values showing extremely good statistical significance (8 orders of magnitude better than is required). The statistics thus testified that the increases were not merely random. The fact that the correct answer was chosen in only 2 out of 10 pre-test questions, but in 9 out of 10 post-test questions, links the learning clearly to the show – especially as there was no improvement evident in the comparison question (Q 3). There was little improvement for the rural students in the original study, who clearly require opportunities to improve the quality of their learning.

Science shows are often viewed merely as providing entertainment and there is little expectation for education to take place. This study suggests that real education is indeed taking place, although again the extent differed dramatically between the different groups. In terms of the 5 key concepts in the show, urban and township students showed dramatic increases in knowledge from pre-test to post-test. Walker *et al*. (2011) found in a similar study (also conducted at Unizulu Science Centre) that there is great potential for improving learning and even changing student attitudes during a science show. Sadler (2004) found both short and long term gains from a similar music show presented in Wales. This provides strong support for the advocacy of science centres in general and science shows in particular when interacting with teachers, government and funders (Wyles, 2010; Kerby et al., 2010; Shaw et al*.,* 2011;).

Table 5.3 in section 5.4 and Figure 5.27 summarise the differences in general performance between the school groups in various areas. In *all* areas the urban student group scored better than the township students who, in turn, scored better than the rural students. This fact obviously needs to be taken into account when planning shows and other activities for each of these groups. I must consider carefully what the factors are which affect this learning for the different groups (Warren et al*.,* 2001). At the end of this study it would appear that English language proficiency is the overriding factor in terms of its influence on the groups and their performance. The implementation of an English language proficiency test confirmed that the limited language of the rural group severely hampered their ability to learn during the show, resulting in their poor performance in the post-test, and poor answers to the written- and other probes.

RQ 3) Building on the data obtained from addressing RQ 2, how can the show be redesigned to improve understanding and minimize conceptual and visualization difficulties with sound?

DBR provided an effective framework for an iterative process to improve science shows which has implications for science shows, interactive exhibits and workshops. It can also be applied to science education and classroom practice. Our modified student difficulties table 5.2 (from student data in the original study), and the CAPS curriculum, provided a good starting point for reviewing the show. Changes were made to misleading slides and new slides were introduced in an attempt to counter difficulties. On the whole these assisted learners but did bring in some new difficulties (for example the “water wave analogy – see section 9.1.2). These changes appeared to assist rural learners substantially as reported in section 9.1.3 and 9.1.4 but did not really appear to assist urban and township groups. This suggests that the two stronger groups may be operating close to the maximum we can expect in terms of learning from one 45-minute show, and that further gains in learning would not really be possible without completely changing presentation style to a far more didactic one. As science centre shows have far more aims than just learning (see the list in section 3.2) it is questionable if this should be our main focus. What was encouraging was seeing that the rural group, who had almost no gains in learning in the original study, now showed some evidence of this. Almost identical pre-test scores (original and subsequent groups) and use of a comparison question allows us to link this improvement to the show itself. They also reported the show as being easier to understand than the rural group in the original study, but were otherwise similar in their answers to attitudinal questions. This makes us confident that the new improved show is more effective for all school groups visiting our centre, as we don’t see the weakest group left behind as much.

The strongest gains were recorded for the group who received instruction in their mother-tongue – isiZulu, in the third iteration of the DBR stages. Despite answering English-language probes, they performed substantially better than the subsequent rural group (drawn from the same school) and even surpassed the township group in terms of their learning gains from pre- to post-test (again matching pre-test scores confirm that gains were made from the show itself). Furthermore a closer look at specific conceptual difficulties (like confusing volume and pitch d 8) showed substantial improvement after the show. A standard English-language proficiency test showed that the gains were not the result of their being better at English. Clearly mother-tongue instruction should be investigated further and means sought to allow students to learn in a way which brings greatest achievement. Greater student confidence was evident in their attitude on the day and their willingness to attempt drawings (section 8.2). Furthermore they responded better to a presenter with whom they clearly identified with more easily than with me, and I felt our centre had made some progress in issues of “cultural congruence” Lee (2003). I feel that this this finding can be added to the literature on the effectiveness of science centres and their programmes (Garnett, 2002).

RQ 4) Building on the data obtained from addressing RQ 1, how can the probes be modified so as to be more effective and valid tools for revealing evidence of student learning and any related conceptual and visual difficulties after experiencing *the sound show*?

As summarised in table 4.4 in section 4.3, the original study produced probes which generally worked well for all groups but less so for the rural group. . Modification of these was informed by the original student data, item analyses and validation exercises. Furthermore an intern group gave insights into the probes which assisted in rewording and reworking them. A summary of changes is given in table 7.1 and these included rewording, replacing weak distractors and misleading drawings. The new distractors performed well (both in the MCQ and the “choose-an-answer” section of the OES) and the new probes showed evidence of improved validity for the subsequent rural group, but similar results (to the original study) for the other two subsequent groups. The original probes had shown good validity for the urban and township groups so it is possible that there was not much room for improvement within the confines of this study. Again it is pleasing to see that the new probes worked better for the rural groups whereas the original probes had shown questionable validity in some areas. We now feel confident that the new probes can be used more effectively across all student groups.

Despite improvements to the open-ended survey it appears that, while the OES yields interesting results (especially of R-M), it is ultimately better suited to a classroom situation where instruction is more systematic and students have time to develop their own concepts and visualizations, than to one forty-five minute show. It is too long and detailed to be performed as a pre- and post-test, so it was difficult to use it as a tool to measure learning in the show and is best suited to looking at student preconceptions. Questions requiring written answers were so poorly handled that the data was often of little use. As mentioned before, student difficulty in translating their mental models into expressed models makes it difficult to analyse learning gains effectively.

Multiple choice questions have many practical advantages in the science centre context and, if well devised, few disadvantages (Abdalla, 2013), and my recommendation would be for these to be the tool of choice in measuring learning gains (R-C) in a science centre show. Interviews provided interesting depth to some of the student drawings but like the OES survey they were time consuming to administer and analyse.

11.3 Limitations of this study

While this study involved quite a few students and many different kinds of data, it was nevertheless limited in many ways. Firstly it was limited geographically, applying only to a very small area and sample of students close to the Science Centre (Richards Bay, KwaZulu-Natal) and, therefore, the findings should not necessarily be generalized to other contexts. Another limitation of the study was that it was a “snapshot” rather than a longer term study, thus, the probing was limited to a single 2-hour interaction with students without follow up or longer-term interventions. Thus, it is not clear whether the learning gains reported have any long term persistence, and this would be an interesting

course for further study. The 6-month interviews showed pleasing evidence of retention, but were not in depth enough to show clear long-term learning gains. The students surveyed were limited in number to 117 (original study),126 (subsequent study) and 42 (final, isiZulu, study) and only Grade 9's were involved. The subject matter was extremely limited, tackling only part of the "sound" section of the syllabus. In addition the study was based on a single science show and therefore there was no control for variations in results that could have been affected by external factors.
Again, while both quantitative and qualitative approaches were used, the research methods were also limited. Multiple choice questions gave only quantitative results and don't explore the reasons for answers or if the answers given were just "lucky". In addition it is possible that students may have learned from the pre-test which would skew the results in the post-test. The issue of "cueing" may have been a factor here and could artificially inflate the post-test score, but this would have affected all three groups in the same way so does not affect our ability to compare them with one another. Drawings were problematic in that students seemed very reluctant to draw and many of these were not even attempted – also some were impossible to interpret. The reasons for this are not clear and warrant further investigation, but some suggestions have been made that student drawings may reflect convenience and confidence-levels, more than misconceptions.. Interpreting written explanations was also not easy and the study confined itself to what could be interpreted directly without further probing. It is clear that the study would have benefited from follow-up Interviews but it was decided at the outset that this aspect fell outside the scope of the study, in particular due to logistical reasons.

11.4 Directions for further research

While the findings of this study has provided a strong motivation for the usefulness of Science Shows as both educational and motivational tools for students, it has also opened up many avenues of potential further study which would add much needed local knowledge to the growing body of (largely foreign) literature in the area of free-choice learning (Rennie *et al*. 2003). These studies would also be of great practical value to science centre staff and teachers. Some possible gaps in our knowledge identified in the above discussions as warranting further research include:

1. Repeating this study at different points in South Africa to see if similar results emerge;
2. Using this data (from 1 above) to compile a comprehensive list of local student difficulties with sound like that reported by Hapkiewicz (1992);
3. Performing a deeper study on naïve theories (as suggested by McCloskey, 1983) in order to:
    a. Characterize these theories in the area of sound, and other areas;
    b. Determine how these naïve theories develop; and,

c. Determine what role these naïve theories play in everyday life.

4. Performing a more detailed study into the effects of language ability on student understanding of sound after attending the show, as well as such effects on student understanding of the probes used at every level of the assessment process (Dempster and Reddy, 2007). Further comparisons could be made between groups experiencing the show in English and in their home language (isiZulu) (Zuma and Dempster, 2008).
5. Performing a detailed study on cultural differences and the effects they have on learning in science, including ways of reasoning, respect for authority figures, questioning, organisational patterns etc. (Lee 2003);
6. Following up this study with more in depth interviews with students, which allow the researcher to probe more deeply into their written answers and especially their drawings;
7. Performing a carefully structured memory test, after a long time-period, to measure retention of different types of demonstrations (like the "CHAMP" analysis of Sadler, 2004);
8. Comparing the effect of the show on local students with students overseas (like my audiences in Beijing, Warsaw and Abu Dhabi) (Falk, 1983);
9. Looking more closely at motivation and enjoyment as factors affecting learning (as per Packer and Ballantyne's, 2002 study);
10. Testing for evidence of the "other intended impacts" of the show listed in 3.2;
11. Designing and evaluating workshops and exhibition materials to complement the show as part of a Science Centre programme (along the lines of Perry and Tisdal, 1994);
12. Applying selected survey questions from the literature to the current study group (i.e. from the studies of Linder, 1993; Huey Por Chang *et al*., 2007; or Caleon and Subramaniam, 2010); and,
13. Continuing with further iterations of the show and probe improvement.
14. Administering another test with a time-lag (say of a week) between show and post-test to ensure that any cueing effect is removed.
15. A comparison of learning gains in science-shows presented pre- and post-instruction in class.
16. A comparison of student understanding and misconceptions between shows with or without use of an oscilloscope to represent sound.

11.5 Implications for Science Centres

11.5.1 Implications for Science Shows

In the future, my goal is that this study will inform and/or provide useful tools for:

1. Those running Science Centres, to monitor and evaluate their programmes and to adapt them to audiences from different educational and cultural backgrounds;
2. Those presenting science shows, to gauge student abilities and to continuously improve their shows and to adapt them to learners from different educational and cultural backgrounds;
3. Teachers of science, to inform them of students' naïve conceptions and mental models and assist them in addressing these; and,
4. Students visiting my Centre, to understand and visualize sound better and to be motivated to follow careers in Science.

The data analysed was extremely revealing in terms of informing best practice in performing science shows. Indeed it pointed out elements in "Good Vibrations" which were actually causing some of the student difficulties, allowing me to modify these. With such clear differences in terms of ability and learning during the show, I need to present shows differently to groups from urban, township and rural schools – especially in the case of rural students. Language is a major issue here (Aikenhead and Jegede, 1999) and translating the show to provide instruction in isiZulu made a major difference. Given the very low pre-test scores, I should perhaps lower my expectations to something more realistic and perhaps even use CPS clickers to perform quick baseline knowledge quizzes before shows. Treagust (2001) showed that learning can be improved with frequent and ongoing assessment of students' understanding. The whole issue of presenting science shows to students who have not covered the work yet in class needs to be carefully rethought. In light of this study it would appear that more would be gained by presenting this show in Grade 10 when students have already done sound in class and when students' conceptions would not be so naïve.

As practitioners, we need to try very hard to match our presentation style and terminology with that of each group (as far as is possible) and to match presenters with audiences. I should consider whether I (with no knowledge of isiZulu) should personally present to rural schools. In this regard, Lee (2003) mentions extensive literature which has indicated discontinuities in the communication and interaction patterns of English-speaking teachers who have students from diverse cultural, language and educational backgrounds, emphasizing the importance of cultural congruence to promote student participation and engagement. The conditions for successful border-crossing (Aikenhead, 1996) should also be considered in the design and presentation of our shows. The gains for the isiZulu group were perhaps linked as much to having a presenter with whom they could identify more than they did with me.

Lee (2003) encourages a balanced orientation to science teaching and learning which "emphasizes academic achievement as well as cultural identity. . . leads students to acquire the language of science as well as their home languages, to understand the culture of science as well as their own cultures, and to behave competently across social contexts." In addition, music has the ability to appeal to our emotions as well as to our brains, and (Lee, 2003) mentions that many students may freely make use of emotion to frame academic arguments. Jansen (2009) stresses the difficulty of changing perceptions in post-apartheid South Africa, and the importance of engaging the emotions:

> "To shift this knowledge in the blood, or understandings of the heart, requires emotional engagement with the subject."

It is beyond the scope of this study to test whether these impacts are happening, but it would make a fascinating topic for further investigation. Finally, I should use the information in the table of student difficulties I produced to inform my shows in terms of content, style, visuals and demonstrations, in order to have maximum impact on students, investigating the appropriate propositional knowledge to build into my shows (Halstead and Anderson, 2009).

11.5.2 Implications for Interactive Exhibits

Acknowledging that shows are limited in time and content, I should devise an exhibition of sound exhibits which could compliment the message of the show, and also allow students a hands-on experience of the phenomenon. In this regard, Allen (2004) describes exhibits as "novel, stimulating, evidence-rich, multi-sensory and fun". Science exhibits provide far more personal choice and allow students to investigate the learning in the show for themselves. Taking into account the differing abilities of the groups, I should ensure that signage and facilitators are prepared to accommodate the different groups. Given my research findings and the varied performance of the different groups, I believe the "APE" (active prolonged engagement) exhibit methodology developed at the San Francisco Exploratorium would be most effective in my context. Studies on these exhibits (Perry and Tisdal, 2004) found that in large part the effectiveness of "APE" exhibits lies in their being designed to be used by more than one member of a social group. This fits in with the group learning favoured by rural schools and would especially benefit that group which needs the most intervention. Students appeared to have problems with diagrammatic reasoning and confusing graphs with pictures, and exhibits and hands on activities can also assist in developing diagrammatic reasoning (Hardy et al. 2005). Thought must be given to signage for the exhibits so that language problems do not hinder their effectiveness. Suggested exhibits could include:

- whisper dishes;
- sound tubes;
- stereo sound tester;
- oscilloscope with microphone to show sound waves;
- different musical instruments;
- electronic instruments;
- interactive models of both transverse and longitudinal waves;
- ringing bell in evacuated bell-jar;
- interactive vocal cords model;
- model of the human ear;
- function generator to test hearing of ultrasound and infrasound and,
- tuning forks.

11.5.3 Implications for Small Group Workshops

Science shows allow the students little or no choice of what to engage in, whereas exploring the exhibits offers them free choice to explore their preferences. Bamberger and Tal (2005) showed that activities which offered *limited choice*, like a small group workshop, offered a scaffolding to the students which helps them develop their natural curiosity into substantial learning. Small group workshops on sound can allow students the chance to discover phenomena working together in small groups, and also to take something home to reinforce the messages learnt. These could be worksheet based, or involve building simple musical instruments from scrap materials, like a drinking straw oboe, clucking cup (from polystyrene cup and string) or various shakers and drums. They could also involve group investigations of wave phenomena, perhaps using slinky springs. Computer-based simulations (like Pasco's "Data Studio" used in the show) could also allow students to experiment with different phenomena, allowing for a greater variety of teaching and learning methods for maximum effectiveness (Calik 2011). The different learning styles of the various groups should be taken into account so that all students are accommodated.

It is my assertion that the mixed-mode model of Science Centres, incorporating the above three activities, works well precisely because it bridges the two extremes of guided instruction and minimal guidance mentioned in the literature review (chapter 2). I believe that different children respond uniquely to different learning experiences and can show different "intelligences" (Gardner, 1985). Twenty years in the science centre have shown me that in any school group there are usually some children who prefer shows, some workshops and some of the more unguided exhibit experience. Having all three modes of instruction during every visit ensures that there will be something for every child to experience.

11.6 Recommendations for science show development and improvement

As a practical consequence of this thesis, I would like to propose the following recommendations for science show development and improvement, based on my findings and using a DBR framework. I have suggested a rough timescale if one has 2 months (40 days) available for this exercise:

Start by selecting a topic for the science show, ideally one around a particular theme (with conceptual coherence) like for example: "magnetism, energy, light" etc. Then perform a literature survey to investigate published research around that topic, looking at content, teaching methodologies and especially student difficulties. Before proceeding, carefully consider the following contexts which will affect your show:

- Your Science Centre context: where is your centre located, what are its aims, what facilities and staff are available for the show?
- The student and school (or visitor) context: who (age, grade, gender, language) will comprise the audience and from what sort of schools do they come (especially socio-economic status, setting – urban or rural etc.)?
- The syllabus context – where does the material you'll be presenting fit into the national curriculum and how can you tie it in to motivate visits by teachers?
- The presenter context – yourself! What skills, experience or "baggage" do you bring to the presentation?

Then, along the lines of appendix a, start to develop your Science Show, preparing a show outline (including which key concepts you want to communicate). Work on a script and PowerPoint slides (if you use this in your shows). Gather the necessary equipment to complement your script and practise using it. Based on what you want to communicate, design effective probes for student evaluation, making sure that questions cover the whole content area and are appropriate for your audiences.

Now start to "field test" your show and probes: perform pilot test of the probes and the show and make minor corrections where necessary. When ready, perform your show under carefully controlled conditions and gather student data. Make sure you perform a validation of probe suitability with the student data before proceeding. Analyse and evaluate student data so that you can draw conclusions on prior knowledge and learning in the show.

Now revise the show in the light of student data for better results, but also revise the probes for better validity. Then transition to the use of an improved model (subsequent show) and improved (subsequent) probes, being sure that you practise the new show before presentation. Perform your subsequent show under carefully controlled conditions and again gather student data using the new probes. Perform a validation of the subsequent probe suitability with student data, to ensure they are still valid and hopefully more valid than before. Analyse and evaluate student data from the subsequent study and draw your conclusions.

Finally consider a targeted revision of the show in the light of subsequent student data to assist the weakest group (or any other group where further change has the potential to make a difference). This may comprise a translation of the show into the group's home language. Redesign subsequent probes if necessary or keep the same probes if they are working well. Transition to the use of an improved model (final show in home language), again making sure you practise new show before presentation). Perform your final show under carefully controlled conditions and gather student data. Perform validation of probe suitability with student data and then analyse and evaluate student data from final study.

Three iterations will usually be sufficient as the process is very time-consuming. At this point you can decide whether further revisions of the show will yield worthwhile gains. This process is summarised in table 11.2 below.

Table 11.2: Recommendations for science show development and improvement

STAGE (FROM FIG. 2.5)		ACTIVITY	TIME (DAYS)	THESIS SECTIONS	SHOW & PROBES
1	Characterization	Select a topic for the science show, ideally one with conceptual coherence	1		*ORIGINAL SHOW* AND PROBES
		Perform a literature survey to investigate published research around that topic	3	Chapter 2	
		Carefully investigate and consider the contexts which will affect your show	1	Section 3.1	
		- Your Science Centre context			
		- The student and school (or visitor) context			
		- The syllabus context – where the material fits into the CAPS curriculum			
		- The presenter context – yourself!			
2	Representation	Develop your Science Show, preparing:		Section 3.2	
		- A show outline (including which key concepts you want to communicate)	1		
		- A script and PowerPoint slides	2		
		- The necessary equipment	2		
		Design effective probes for student evaluation	2	Sect. 3.3, 3.4 & 3.5	
		Perform pilot test of probes and show. Make minor corrections where necessary	3	Sect. 3.4.1 & 3.4.2	
3	Field Test	Perform your show under carefully controlled conditions and gather student data	1		
		Perform validation of probe suitability with student data.	2	Chapter 4	
		Analyse and evaluate student data	2 (20)	Chapter 5	
4	Redesign	Revise the show in the light of student data for better results	2	Chapter 6	*SUBSEQUENT* SHOW & PROBES
		Revise the probes in the light of student data for better validity	2	Chapter 7	
2 (2nd)	Representation (2nd iteration)	Transition to the use of an improved model (subsequent show) & improved (subseq.) probes	1	(As above)	
		(Practise new show before presentation)			
3 (2nd)	Field Test (2nd iteration)	Perform your subsequent show under carefully controlled conditions and gather student data	1		
		Perform validation of subsequent probe suitability with student data.	2	Chapter 8	
		Analyse and evaluate student data from subsequent study	2	Chapter 9	
4 (2nd)	Redesign (2nd iteration)	Targeted revision of the show in the light of subsequent student data to assist weakest group	2	Chapter 10	*FINAL SHOW* (TRANSLATION?)
		(Possibly translation into group's home language?)		Sect. 10.1	
		Redesign subsequent probes if necessary or keep the same probes	(0)		
2 (3rd)	Representation (3rd iteration)	Transition to the use of an improved model (*final* show in home language?)	1	(As above)	
		(Practise new show before presentation)			
3 (3rd)	Field Test (3rd iteration)	Perform your final show under carefully controlled conditions and gather student data	1	Sect. 10.2	
		Perform validation of probe suitability with student data if necessary	2	Sect. 10.3	
		Analyse and evaluate student data from final study	2		
4 (3rd)	Further Redesign (3rd iteration)	Do final revisions and decide whether further revisions of the show will yield worthwhile gains.	2 (40 DAYS)	Sect. 11.4 Future study	?

11.7 Conclusion

The dire backdrop painted in chapter 1 of education in SA should serve as both a warning and a motivation. The "border-crossing" (Aikenhead, 1996) which is necessary for our students to accommodate new scientific ideas is not going to happen in unresourced schools with unqualified teachers. We all need to provide opportunities for learning to take place with appropriate demonstrations and hands-on activities. As aptly stated by Jonathan Jansen:
"We must push back against mediocrity. We must measure our success not by the results of the students who pass well, but by the results of the hundreds of thousands who fail and pass poorly every year." (Jansen, 2012)

In education, we all need to make allowance for the very different contexts in which our students learn, and provide effective instructional scaffolding:

> "For these students, effective instructional scaffolding takes into account students' linguistic and cultural backgrounds as well as science experience. The aim is to encourage students to question and inquire without devaluing the norms and practices of their homes and communities. Through progressing along the teacher-explicit to student-exploratory continuum, teachers encourage students to take initiative and assume responsibility for their own learning." (Lee, 2003, pg. 477).

Science Centres are well positioned to assist in all these processes.

Many studies (see 2.5) show that science centres have far reaching impacts and can be shown to have an important impact on the science literacy of surrounding communities (Falk and Needham 2011) . Furthermore science enrichment experiences involving novel concepts (like a visit to a science centre) have the flexibility to present concepts differently from how they are being presented in the classroom (Caleon and Subramaniam (2007). It is my *belief* that science centres have an increasingly important role to play in assisting and augmenting our education system. It is my *hope* that this study may serve to inform and motivate more people in the science centre field. In the words of the great physicist, Richard Feynman:

> *"Study hard what interests you the most in the most undisciplined, irreverent and original manner possible."*

References

Abdalla, M. E. (2013). “Multiple Choice Questions Revisited: Improvement of Validity for Fair Tests.” Gezira Journal of Health Sciences, ***9*****(1).**

Abell, S. K., Smith, D. C. (1992). "What is science? Preservice elementary teachers' conceptions of the nature of science." The history and philosophy of science in Science Education. Proceedings of the international conference on the history and philosophy of science and science teaching. **Volume I**. S. Hills. Kingston, Ontario, The Faculty of Education, Queens University**:** 11-22.

Aikenhead, G. (1996). "Science Education: Border crossing into the subculture of science." Studies in Science Education **27**: 1-52.

Aikenhead, G. S., and Jegede, 0. J. (1999). "Cross-cultural science education: A cognitive explanation of a cultural phenomenon." Journal of Research in Science Teaching, **36**(3), 269-287.

Allen, S. (2002). "Looking for learning in visitor talk: a methodological exploration." Learning Conversations in Museums (eds. Leinhardt G, Crowley K and Knutson K) published by Lawrence Erlbaum Assoc. Mahwah NJ: 259-303.

Allie, S., Buffler, A., Kaunda, L., Campbell, B., Lubben, F. (1998). "First-year physics students' peceptions of the quality of experimental measurements." International Journal of Science Education, **20**(4), 447-459

Anderson, D., Lucas, K. B. , Ginns, I. S. , Dierking, L. D. (2000). "Development of knowledge about electricity and magnetism during a visit to a science museum and related post-visit activities." Science Education **84**(5): 658-679.

Anderson, T.R., Schönborn, K. J., du Plessis, L. Gupthar, A.S. and. Hull, T. L (2013). "Identifying and Developing Students’ Ability to Reason with Concepts and Representations in Biology (Ch. 2)" in Treagust, David F.; Tsui, Chi-Yan (Eds.) Multiple Representations in Biological Education, Series: Models and Modeling in Science Education, Vol. 7 pp 19-38, The Netherlands: Springer.

Bamberger, Y., Tal, T. (2006) "Learning in a Personal Context: Levels of Choice in a Free Choice Learning Environment in Science and Natural History Museums." Science learning in everyday life. (Lynn D. Dierking and John H. Falk Section Coeditors)

Barriault, C. (2001). "The Science Center Learning Experience: A Visitor-Based Framework." The Informal Learning Review.

Boyes, E., and Stanisstreet, M. (1991). "Development of pupils’ ideas about seeing and hearing—The path of light and sound." Research in Science and Technology Education*,* **9***,* 223–251.

Brenner, E. (2010). "Interactive teaching using clickers." Education South Africa(Feb): 28.

Bradburne, J. M. (1998). “Dinosaurs and white elephants: The science center in the twenty-first century.” Public Understanding of Science, 7(3), 237-253.

Brown, A. L. (1992). “Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings.” The journal of the learning sciences, ***2*****(2),** 141-178.

Caleon, I., Subramaniam, R. (2005). "The impact of a cryogenics-based enrichment programmes on attitude towards science and the learning of science concepts." International Journal of Science Education **27**(6): 679-704.

Caleon, I., Subramaniam, R. (2007). " Augmenting Learning in an Out-of-school Context: The Cognitive and Affective Impact of Two Cryogenics-based Enrichment Programmes on Upper Primary Students." Research in Science Education (2007) **37**:333–351

Caleon and Subramaniam, (2010). "Development and Application of a Three-Tier Diagniostic test to Asses Secondary Students' Understanding of Waves." International Journal of Science Education, **32***(7*), 939-961.

Calik, M et al. (2011). "A Comparison of Different Conceptual Change Pedagogies Employed Within the Topic of 'Sound Propagation'.'' Journal of Science Education and Technology (2011) **20**:729–742

Chi, M. T. H., and Roscoe, R. D. (2002). "The process and challenges of conceptual change." In M. Limon and L. Mason (Eds.), "Reconsidering conceptual change: Issues in theory and practice" (pp. 3-27). Dordrecht: Kluwer.

Christie, P (1986). "The Right to Learn": Sached Ravan : pg. 12

Cobern, W. W., Aikenhead, G. S. (1998). "Cultural aspects of learning science." International handbook of Science Education, Part 1. B. J. Fraser, Tobin, K. G. Dordrecht, Netherlands, Kluwer Academic Press**:** 39-52.

Collins, A. (1992). "Toward a design science of education." New directions in educational technology (pp. 15-22). Springer Berlin Heidelberg.

Cooke, N. J., & Breedin, S. D. (1994). "Constructing naive theories of motion on the fly." Memory & Cognition, ***22*****(4)**, 474-493.

Crouch, C., Fagen, A.P., Callan, J.P. and Mazur, E., (2004). "Classroom demonstrations: Learning tools or entertainment?". American journal of physics, ***72*****(6)**, 835-838.

Dempster, E. (2004) "Summary report on the evaluation of the Senior certificate examination for Umalusi."

Dempster, E. R., & Reddy, V. (2007). "Item readability and science achievement in TIMSS 2003 in South Africa." Science Education, ***91*****(6)**, 906-925.

Deregowski, J.B. (1989). "Real space and represented space: Cross-cultural perspectives." Behavioural and Brain Sciences (1989) **12**: 51 - 119

DiSessa, A. A. (2002). "Why conceptual ecology is a good idea." In M. Limon and L. Mason (Eds.), "Reconsidering conceptual change: Issues in theory and practice" (pp. 29-60). Dordrecht: Kluwer.

Disessa, A. A. (1988). "Knowledge in pieces." 49 - 70

Driver, R. (1989). "Students' conceptions and the learning of science." International Journal of Science Education **11**: 481-490.

Duit, R., Treagust, D. F. (2003). "Conceptual change: A powerful framework for improving science teaching and learning." International Journal of Science Education **25**(6): 671-688.

Edelson, D. C. (2002). “Design research: What we learn when we engage in design.” The Journal of the Learning sciences, ***11(1),*** 105-121.

Eisenhart, M. (2001). "Educational ethnography past, present, and future: Ideas to think with." Educational Researcher **30**(8): 16-27.

Eshach, H., and Schwartz, J. L. (2006). "Sound stuff? Naive materialism in middle-school students' conceptions of sound." International Journal of Science Education, **28**(7), 733-764 // g6, P, S, GC.

Falk, J. (1983). "A cross-cultural investigation of the novel field trip phenomenon": National Museum of Natural History, New Delhi.

Falk, J. H., Dierking, L.D. (1992). "The Museum Experience." Whaleback Books.

Fallk, J. H., Dierking, L.D. (2000). "Learning from Museums." Altamria Press.

Falk, J. (2009). "Identity and the Museum Visitor Experience." Left Coast Press

Falk, J. H., Gillespie, K.L. (2009). "Investigating the Role of Emotion in Science Center Visitor Learning." Visitor Studies **12**(2): 112-132.

Garnett, R. (2002). "The Impact of Science Centers/Museums on their Surrounding Communities: Summary Report." ASTC Website: www.astc.org.

Gilbert, J.K., Boulter, C. and Rutherford, M. (1998). "Models in explanations, Part 2: Whose voice? Whose ears?" International journal of science education **20**(2): 187-203.

Giroux, H. (1992). Border crossings: Cultural workers and the politics of education. New York: Routledge.

Gouniyal, N. (2007). "Science Theatre: The Nature of Learning An evaluation of visitor learning from Science Theatre shows at Techniquest, Cardiff." Unpublished Masters Thesis.

Graue M. E. (1993) "Integrating theory and practice through instructional assessment." Educational Assessment **1**, 283–309.

Grayson, J., Anderson, T. R. , Crossley, L. G. (2001). "A four-level framework for identifying and classifying student conceptual and reasoning difficulties." International Journal of Science Education **23**(6): 611-622.

Hanushek, E. A., & Wößmann, L. (2007). “The role of education quality for economic growth.” World Bank Policy Research Working Paper, (4122)**.**

Halstead, S.E. and Anderson, T.R. (2009) "What can we learn from studies on students' conceptual difficulties? The case of propositional knowledge in acid-base models. "In Schafer, M and McNamara, C. Proceedings of the Seventeen Annual Meeting of the Southern African Association for Research in Mathematics, Science and Technology Education. Vol 1: Long papers (pp 82-93) Rhodes University, Grahamstown. (ISBN # 978-92-990043-6-4)

Hapkiewicz, A. (1992). "Finding a List of Science Misconceptions." MSTA Newsletter, 38(Winter'92), pp.11-14.)

Hardy, I., Schneider, M., Jonen, A., Stern, E., Moller, K. "Fostering diagrammatic reasoning in science education." Swiss Journal of Psychology **64** (3), 2005, 207–217
Harvey, S. (1999). "Phasing science InSET in developing countries: reflections on the experience of the Primary Science Programme in South Africa." International Journal of Science Education, 21(6), 595-609.

Houle and Barnett (2008). "Students' Conceptions of Sound Waves Resulting from the Enactment of a New Technology-Enhanced Inquiry-Based Curriculum on Urban Bird Communication." Journal of Science Education and Technology *(2008)* **17***:242-251*

Hrepic et al. (2010). "Identifying students' mental models of sound propagation: The role of conceptual blending in understanding conceptual change." Physics Education Research **6**, 020114 _2010_

Huey-Por Chang et al. (2007): "Investigating Primary and Secondary Students' Learning of Physics Concepts in Taiwan." International Journal of Science Education, **29**(4), 465-482

James, A., Naidoo, J., & Benson, H. (2008). "CASME's approach to the sustainability of science education in south Africa". In XIII. IOSTE Symposium, the use of science and technology education for peace and sustainable development (pp. 21-26).

Jansen, Jonathan D.(1998) "Curriculum reform in South Africa : a critical analysis of outcomes-based education". Open UP: Research Articles of the University of Pretoria.

Jansen, J. (2008). "Opening the doors of learning: Changing schools in South Africa." Perspectives in Education **26**(4): 115-116.

Jansen, J. D. (2009). "When Politics and Emotion Meet Educational Change In Racially Divided Communities." Phi Delta Kappan **90**(5): 327-332.

Jansen, J. (2012). "National Senior Certificate results belie conceptual and skill limitations of school-leavers." South African Journal of Science **108**(3-4): 9-10.

Kerby (2010) "Fusion Science Theater Presents The Amazing Chemical Circus: A New Model of outreach That Uses Theater To Engage Children in Learning." Journal of Chemical Education _ **87** (10) October 2010

Kirschner P, Sweller J, Clark R. "Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching." Educational Psychologist, **41***(2), 75 – 86*

Kose, S. (2008). "Diagnosing Student Misconceptions: Using Drawings as a Research Method." World Applied Sciences Journal **3**(2): 283-293.

Lagemann, E. C. (2002). "An elusive science: The troubling history of education research." University of Chicago Press.

Lee, O. (1999). "Science knowledge, world views, and information sources in social and cultural contexts: making sense after a natural disaster." American Educational Research Journal **36**(2): 187-219.

Lee, O., and Fradd, S. H (1998). "Science for all, including students from non-English language backgrounds." Educational Researcher **27**(3): 12-21.

Lee, O. (2003). "Equity for Linguistically and Culturally Diverse Students in Science Education: A Research Agenda." Teachers College Record **105**(3): 465-489.

Lee O, Luykx A, Buxton C and Shaver A, (2007) "The Challenge of Altering Elementary School Teachers' Beliefs and Practices Regarding Linguistic and Cultural Diversity in Science Instruction" Journal Of Research In Science Teaching **44** (9), 1269–1291

Levin, J. R., and O'Donnell, A. M. (1999). "What to do about educational research's credibility gaps?" Issues in education, ***5*****(2)**, 177-229.

Linder, C. J., Erickson, G. L. (1989). "A study of tertiary physics students' conceptualizations of sound." International Journal of Science Educatio*n,* **11**, 491-501

Linder, C. J. (1992). "Understanding sound: so what is the problem?"Physics Education*,* **27**(5),258-264

Linder, C. J. (1993). "University physics students' conceptualizations of factors affecting the speed of sound propagation." International Journal of Science Education*,* **15**(6), 655-662

Linder, C. J. (1993). "A Challenge to Conceptual Change." Science Education **77**(3): 293-300

Matthews, M. R. (1993). "Constructivism and Science Education: Some epistemological problems." Journal of Science Education and Technology **2**(1): 359-370.

McCloskey, M. (1983). "Naive theories of motion." Mental models, 299-324.

Mazur, E. (2009). "Farewell, lecture?" Science **323**(January): 50, 51.

Naidoo, P., & Lewin, K. M. (1998). "Policy and planning of physical science education in South Africa: Myths and realities." Journal of research in science teaching, *35*(7), 729-744.

Niebert, K., S. Marsch, et al. (2012). "Understanding needs embodiment: A theory-guided reanalysis of the role of metaphors and analogies in understanding science." Science Education **96**(5): 849-877.

Packer, J., Ballantyne, R. (2002) " Motivational factors and the visitor experience: A comparison of three sites." Curator, **45** (3), 183-198

Penfield, R.D., Lee, O. (2009) "Test-Based Accountability: Potential Benefits and Pitfalls of Science Assessment With Student Diversity." Journal Of Research In Science Teaching. **47** (1), 6–24 (2010)

Perry, D., and Tisdal, C. (2004). "*Going APE! at the Exploratorium: Interim Summative Evaluation Report".* (Available from the Exploratorium, 3601 Lyon St, San Francisco, CA 94123).

Persson, P. E. (2000). "Science centers are thriving and going strong!" Public understanding of science, 9(4), 449-460

Pfundt, H and Duit, R (1988) "Bibliography: Students‘ Alternative Frameworks and Science Education." Kiel. IPN

Phelan, P., Davidson, A., and Cao, H. (1991). "Students' multiple worlds: Negotiating the boundaries of family, peer, and school cultures." Anthropology and Education Quarterly, **22**, 224-250.

Phillips, D. C., & Dolle, J. R. (2006). “From Plato to Brown and beyond: Theory, practice, and the promise of design experiments.” Instructional psychology: Past, present and future trends. Sixteen essays in honour of Erik De Corte, 277-292.

Pines, L., West, L. (1986). "Conceptual understanding and science learning: An interpretation of research within a sources of knowledge framework." Science Education **70**(5): 583-604.

Posner, G. J., Strike, K. A. , Hewson, P. W. , Gertzog, W. A. (1982). "Accommodation of a scientific conception: Toward a theory of conceptual change." Science Education **66**(2): 211-227.

Prediger, S., Gravemeijer, K., & Confrey, J. (2015). “Design research with a focus on learning processes: an overview on achievements and challenges.” ZDM, ***47*****(6),** 877-891.

Reddy, V. (2006). “Mathematics and science achievement at South African schools in TIMSS 2003.” HSRC Press.

Reddy, V., Prinsloo, C., Visser, M., Arends, F., Winnaar, L., & Rogers, S. (2012).” Highlights from TIMSS 2011: the South African perspective.” HSRC, Pretoria.

Reeves, T. C. (2006). “Design research from a technology perspective.” Educational design research, ***1*****(3),** 52-66.

Reiner, M., Slotta, J. D., Chi, M. T. H., and Resnick, L. B. (2000). "Naïve physics reasoning: A commitment to substance-based conceptions." Cognition and Instruction **18**(1), 1–34.

Rennie, L., Feher, E., Dierking, L., Falk, J. "Towards an agenda for advancing research on science learning in out of school settings." Journal Of Research In Science Teaching **40** (2), 112–120 (2003)

Rodriguez, A. J. (1997). "The dangerous discourse of invisibility: A critique of the NRC's National Science Education Standards." Journal of Research in Science Teaching **34**: 19-37.

Rodriguez, A. J. (1998). "Busting open the meritocracy myth: Rethinking equity and student achievement in science education." Journal of Women and Minorities in Science and Engineering **4**(2, 3): 195-216.

Sadler, W. J. (2004). "Evaluating the short and long-term impact of an interactive science show." Unpublished Masters Thesis.

Salmi, H. (2003). "Science Centres as Learning Laboratories: experiences of Heureka, the Finnish Science Centre." International Journal of Technology Management **25**(5): 460-476.

Schönborn, K.J., Anderson, T.R., and Grayson, D.J. (2002). Student difficulties with the interpretation of a textbook diagram of immunoglobulin G (IgG). Biochemistry and Molecular Biology Education, **30**, 93–97.

Schönborn, K.J. and Anderson, T.R. (2008) "Bridging the Educational Research-Teaching Practice Gap: Conceptual understanding, Part 2: Assessing and developing student knowledge." **Biochemistry and Molecular Biology Education** **36** (5), 372-379.

Schönborn, K.J. and Trevor R. Anderson (2009): "A Model of Factors Determining Students' Ability to Interpret External Representations in Biochemistry." International Journal of Science Education, **31**:2, 193-232

Segall, M. H., D. T. Campbell, and M. J. Herskovits (1966). "The influence of culture on visual perception." Bobbs-Merrill." New York (1966).

Shaw, A. J., Harrison, T. G., Shallcross, K. L., Williams, S.J. and Shallcross, D.E. (2011). "On the impact of the Bristol ChemLabS outreach programme on admissions to the School of Chemistry." New Directions in the Teaching of Physical Sciences **7** : 22-26.

Sjoberg, S., (2007) "Constructivism and learning." *Invited contribution to Baker, E.; McGaw, B. and Peterson P (Eds) (2007)* International Encyclopaedia of Education 3rd Edition, *Oxford: Elsevier (in print)*

Solano-Flores, G., and Nelson-Barber, S. (2001). "On the cultural validity of science assessments." Journal of Research in Science Teaching, **38**, 553–573.

Stocklmayer, S. and Gilbert, J.K., (2002). "New experiences and old knowledge: towards a model for the personal awareness of science and technology." International Journal of Science Education, **24(8):** 835-858.

Stocklmayer, S. M., Gore, M.M. and Bryant, C. (2001). "Science Communication in Theory and Practice." Kluwer Academic Publishers: 219-235.

Stokes, D. E. (1997). "Pasteur's Quadrant: Basic Science and Technological Innovation." Brookings Institution Press. Washington (DC).

Thomas, G.V. and A.M.J. Silk, 1990. "An introduction to the psychology of children's drawings." Hemel Hempstead: Harvester Wheat Sheaf.

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

Trujillo, C. M., Anderson, T. R., & Pelaez, N. J. (2016). "An instructional design process based on expert knowledge for teaching students how mechanisms are explained." Advances in physiology education, ***40*****(2),** 265-273.

Van Fraassen, B. C. (2010). "Scientific representation: Paradoxes of perspective." Analysis, ***70*****(3),** 511-514.

Von Aufschnaiter, C. and Von Aufschnaiter, S. 2007. "University students' activities, thinking and learning during laboratory work." European Journal of Physics **28,** S51-S60.

von Glaserfeld, E.(2010) "Why people dislike radical constructivism." Constructivist Foundations **6** (1)

Walker, G., Stocklmayer, S.M., Grant, W.J. (2011). "Science Theatre: Changing South African Students' Intended Behaviour Towards HIV AIDS." International Journal of Science Education

Wang, F., & Hannafin, M. J. (2005).” Design-based research and technology-enhanced learning environments.” Educational technology research and development, ***53*****(4),** 5-23.

Warren, B., Ballenger, C., Ogonowski, M., Roseberyy, A., Hudicort-Barnes (2001) "Rethinking Diversity in Learning Science: The Logic of Everyday Sense-Making" Journal Of Research In Science Teaching **38**, (5), 529-552 (2001)

Watt, D., and Russell, T. (1990). "Sound, primary SPACE project research report." Liverpool, England: Liverpool University Press.

Wittmann, M. C., Steinberg, R. N., Redish, E. F. (2003). "Understanding and affecting student reasoning about sound waves." International Journal of Science Education, **25**(8), 991-1013

Zuma, S.C. and Dempster, E.R. (2008). “isiZulu as a language of assessment in science.” African Journal of Research in Mathematics, Science and Technology Education, 12: pp.31–46.

Paper published from this thesis

Appendices

Appendices (a) to (e) are included in this thesis and follow this page.

Appendices (f) to (p) are included on the electronic copy on the CD as auxiliary material.