Using Adapted Primary Literature to Test the Understanding of Concepts of Evidence in Chemistry held by First Year University Undergraduate Students

by

ELIZABETH VERGIS

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Theoretical, Cultural and International Studies in Education

Department of Educational Policy Studies
University of Alberta

ABSTRACT

Adapted primary literature (APL) is a genre which still retains the canonical form and authentic results of primary literature, but is made readable and understandable for a target population which in this study consists of first-year university undergraduate students. The rationale underlying such adaptations is that they are more consistent with the nature of scientific inquiry than undergraduate chemistry textbooks. Learning facilitated by the usage of primary literature may be a way of developing the capacity for scientific ways of thinking among undergraduate students, as demonstrated by the pioneers in the APL field, Baram-Tsabari & Yarden. My objective in this study was to use APL in order to probe the understanding of Concepts of Evidence held by first-year undergraduate students; and investigate whether and how a pedagogic intervention could develop this understanding further. Concepts of Evidence are ideas that dictate how evidence can be collected, verified, analyzed, and interpreted. In order to achieve my goal I had to: (a) identify one piece of primary literature (b) compose an APL based on this primary literature; (c) devise a measure of Concepts of Evidence called the "Evidence Survey," comprised of a pre-test questionnaire and a post-test questionnaire; (d) identify college/university classes where the instructors teaching first-year chemistry courses read the APL and agreed to be interviewed; (e) enlist students from these classes who were willing to participate in this study; (f) distribute the APL to these students followed by the administration of the pre-test questionnaire; (g) render a pedagogic intervention to the participating students on the Concepts of Evidence followed by the post-test questionnaire; (h) analyze the data from the two questionnaires to make the desired comparisons; and (i) transcribe the interviews and code the transcripts. Thematic content analysis was carried out on the coded transcripts and the main themes that emerged were identified. Four themes emerged from the student interviews and five from interviewing the instructors. My findings from the Evidence Survey showed that the majority (86%) of the student participants performed better on the post-test, suggesting that the teaching intervention was effective in furthering the understanding of both the content knowledge of the APL and the Concepts of Evidence embedded in it.

Preface

This thesis is an original work by Elizabeth Vergis. The research project, of which this thesis is a part, received research ethics approval from the University of Alberta Research Ethics Board. Project name "Using Adapted Primary Literature to teach and test the understanding of Concepts of Evidence in chemistry held by high school and university undergraduate students."

No. Pro00012614, February 22, 2010

Dedication

To my Ammachy (Mom) and Appachen (Dad)

with heartfelt gratitude for their unconditional love and support

Acknowledgements

I can do all things through Christ who strengthens me.
(Philippians 4:13)

Above all else, I must thank my Heavenly Father for providing me this opportunity to do research and strengthening me along the journey. He provided me with the grace to run this race and overcome all obstacles that arose along the way. Through Him all things are possible.

There have been many individuals who have helped me as I journeyed through my doctoral program, and to these people I want to give thanks. I cannot possibly name all of you, but know that you have made a difference.

My thanks go out to, all the students and Instructors of Chemistry 101 or its equivalent, who took part in my study. I feel humbled and honoured by the trust you placed in me and of the way you gave of your time and effort to be part of my study. A special thank you to the Instructors for generously contributing their time, knowledge and wisdom during their interviews.

Thank you to Dr. Randy Wimmer, my Doctoral Supervisor, for your interest, your time, your support and guidance which were invaluable. The opportunities to work alongside you and to learn with and from you in my roles as Doctoral student, research assistant and colleague have been greatly valued. You are a great mentor and teacher educator.

Thank you to my Supervisory Committee, Dr. André Grace and Dr. José da Costa, who challenged my thinking through their provocative questions and comprehensive feedback. I have valued your time and expertise.

Thank you to my Examining Committee members who challenged me to go even further in my thinking by posing thought-provoking questions. Thank you – Drs. Evelyn Steinhauer and Norma Nocente. Thank you also to Dr. Bonnie Stelmach for agreeing to chair my defense.

Special thanks to my external examiner, Dr. Alison Flynn, whose research is so inspiring.

I am also grateful to Dr. Glen Loppnow ho allowed me to pilot my Evidence Survey and conduct interviews in his Science 100 class. He also thoroughly checked my APL as it was being adapted from the primary literature.

Thank you to Dr. Margaret-Ann Armour who listened to me patiently and encouraged me to always perform at my best. Special thanks go out to you for giving so freely of your time, knowledge and wisdom. Thank you for being a role model for me.

Thank you to those family and friends who continue to support me in both my academic and personal lives. Your unwavering confidence, endless patience, gentle encouragement, and ongoing cheerleading have helped me throughout this journey. Above all others, I must thank my parents. If I have been successful, it is because you are and have always been so exceptional and supportive. Your fervent prayers and constant encouragement are greatly cherished. Thank you – Mom and Dad (Gracey and Mathew Daniel).

Thank you to my beloved husband and daughter who are my greatest supporters. The way you two have blessed my life is hard to capture in words. Know that you are my inspiration and greatest source of love and joy. I could always count on you in my times of need. Thank you, Mathew and Priyanka. Always and for ever...

Finally, I would like to express my gratitude for the financial support provided by the National Science and Engineering Research Council, the University of Alberta (Profiling Queen Elizabeth II Award two times, Cathy MacPhail Career Development Scholarship, Dr. Peggy Quinney Doctoral Scholarship and Myer Horowitz Graduate Student Award) the Department of Educational Policy Studies (Educational Policy Graduate Students Scholarly Conference Award), the Graduate Students' Association (Graduate Students' Scholarly Conference Award) and the Centre for Research in Youth, Science Teaching and Learning (Graduate Scholarly Conference Award).

TABLE OF CONTENTS

Page

CHAPTER I. INTRODUCTION

CHAPTER II. LITERATURE REVIEW

CHAPTER III. RESEARCH DESIGN: METHODS AND METHODOLOGY

CHAPTER IV. FINDINGS OF THE STUDY

CHAPTER V. DISCUSSION OF FINDINGS

List of Tables

List of Figures

Page

List of Appendices

Pages

CHAPTER ONE

INTRODUCTION

Prelude: The Opening

The sun was slowly rising in Aluva, a town in the state of Kerala, South India. The azure sky was a mixture of yellow, orange, and crimson. A rooster somewhere was crowing loudly "Cock-a-doodle-doo." In the distance one could also hear the cry of a peacock, strutting with its feathers fully spread out, wooing a peahen nearby. The peahen was crying very sweetly, sounding like raindrops. The paddy fields nearby were flanked by tall coconut palms that towered over them majestically. In the early hours of dawn, the paddy fields lay peaceful awaiting awakening.

From where I was standing, I could see a number of kitchen lamps lit in the surrounding houses. Most of the women in these houses were brewing their first cups of coffee sweetened with saccharin. They would then serve coffee to every adult member of the house and then drink up their own cup of coffee. The smell of freshly-brewed coffee wafted into the fresh morning air. There was a hustle and an air of expectation as women changed and got ready to sow the paddy fields with rice. These paddy fields were divided into small, rectangular or square segments separated by partitions called *varambu* in the local language. The fields were ploughed, fertilized, and waiting to receive the seeds.

The rice seeds were made ready by dipping them in water mixed with cow-dung, and keeping them in the same wet sack for one day. The seeds sprouted by the next day. The women came down from the houses dressed in colourful sarongs or *mundu* as it is called in Kerala, and *chattas*, a V-necked white blouse. They were wearing a head-cloth or a small umbrella on their heads in order to protect them from the scorching heat of the sun. The ladies walked onto the *varambu*, the partitions to the paddy fields, and were given baskets containing the sprouting rice grain.

Each woman with a basket in hand walked down the *varambu* to the segment allotted to her. They took handfuls of the sprouted seeds from their baskets and threw them into the specially ploughed area in each segment of the paddy field. Once the ladies had thrown all the seeds, they handed back their baskets, then chatted among themselves and walked back to their houses. The sprouting grains have to be left alone, entrusted to Mother Nature for a couple of weeks, to grow into delicate rice seedlings which have to be watered and watched.

To supplement Mother Nature's helping hand, the Kerala Government has adopted information technology for delivering information and services to all farmers, especially the rice-farmers. The advantage of having this information technology is that it helps to deliver the right information, to the right people, at the right time, and in the right context. Technological advancements have thereby enabled farmers to avert disasters brought on by the weather. Information technology has advanced not only in India, but has progressed all over the world, as a result of the rampant globalization of this century.

I start this work by using the above metaphor about rice farming in the town of Aluva. Each stage in the cultivation of rice will be used as a symbol in each of the five chapters of my thesis. This metaphor uses something that can be observed, such as the stages in rice cultivation, to convey the essence of each chapter, which is abstract. This vivid comparison is apt because it illustrates graphically the main function of each chapter. I will return to the rice farming metaphor throughout this work.

Background

The 21st century has been host to a tremendous acceleration in technological progress. Within the next 30 years it may be possible to create superhuman intelligence referred to as the dawn of *technological singularity* (Vinge, 2012, 1993). According to Vinge, science may attain technological singularity whereby users will be superhumanly intelligent due to the interaction with computers.

These giant leaps made by technology have enabled people to learn and teach through mediums like Twitter, Instagram, and e-Class. This access to multimedia has influenced pedagogy and program delivery in Canada, where a new mode of teaching has emerged in the form of online learning. In the present climate, effective teaching requires extensive use of technology. Instructors have to cease thinking about technology as a supplemental teaching tool and envision it as essential for successful student learning (Ertmer & Ottenbreit-Leftwich, 2010).

When faced with the prospect of such stupendous, scientific breakthroughs, science educators should, I think, join Durant (1994) in asking: “What is it reasonable to hope and expect that ordinary citizens will know about science in order to equip them for life in a scientifically and technologically complex culture?” (p. 83). One cannot answer Durant’s question without emphasizing the need to have a citizenry that is scientifically literate (Laugksch, 2000; Linder, Östman, & Wickman, 2007; Millar, 2006; Roth, 2003; Ryder, 2001). Although debates about what is meant by a functional scientific literacy have been entertained since the 19th century, there is still controversy about appropriate definitions. In the meantime, there have been many attempts to generate a school science curriculum whose chief objective is to equip students for living successfully in a world that is scientifically and technologically complex (Linder, Östman, & Wickman, 2007; Millar, 2006; Roth, 2003; Ryder, 2001). The present state of school science, however, leaves much to be desired (Rosenblatt, 2011; Marshall, 2003; Millar & Osborne, 1998).

The State of School Science

The state of science education in schools has been the topic of much discussion during the last two decades (Marshall, 2003; Millar, 1996; Millar & Osborne, 1998; Osborne, Driver & Simon, 1998; Weiss, Banilower, McMahon, & Smith, 2001), because there were two issues of particular concern to science educators:

1. Pupils' attitudes towards science as a school subject;
2. Pupils' understanding of the ideas of science.

A strong indicator of the pupils' attitude is their choice of subjects. The proportion of post-secondary students taking science, particularly physical science, remains a cause for concern (Canadian Council on Learning, 2007; Osborne, Driver, & Simon, 1998). "For whatever reasons, large numbers of pupils, particularly girls, are rejecting science as soon as they get the chance" (Gott & Johnson, 1999, p. 21; Kumar & Morris 2005; Wiens, Depping, Wallerich, Van Laar & Juhl, 2003). "Girls are evenly represented in biology and outnumber boys in chemistry, but are underrepresented in physics" (Office for Civil Rights (ED), 2012, p. 2; National Science Foundation, 2008).

The International Mathematics and Science Study (TIMSS) is a large-scale study conducted by the International Association for the Evaluations of Educational Achievement (IEA). TIMSS provides information about how Canadian elementary and secondary students in various grades understand ideas of science and perform on international mathematics and science tests. These studies indicate that students' interest and performance in mathematics and science deteriorate through elementary and secondary school, and that their participation declines in high school (Bordt, De Broucker, Read, Harris, & Zhang, 2001). Particularly, "interest in mathematics and science declines between Grade 4 and Grade 8 and continues to drop during high school" (p. 9). In 1995, only 42% of students were taking both these courses in their last year of high school (Bordt et al., 2001). Many students find mathematics and science "difficult" or "boring" (p. 9). Students who do not believe that they will be successful in science are likely to avoid science-related activities and invest less time for them (Britner & Pajares, 2006). "Even when they have done well in mathematics and science in the past, and believe that these subjects are important to them if they want to succeed in life*,* many students are unwilling to pursue them" (Bordt et al., 2001, p. 9). Perhaps we should pause to consider what could be the root cause of this student unwillingness to invest in science?

The science taught in our schools has not changed drastically and essentially remains the same as it was about 50 years ago (Marshall, 2003). The current science education reform movement in the United States recognizes that the traditional goals of science education are out of date and new curricula need to be "invented" to replace them (Hurd, 2000, p. 282). However, we are living at present, in a "knowledge-intensive era" (p. 282), and the Internet enables students to have access to nearly all that is known in the various disciplines of science. Due to the immense changes occurring in science and technology, in society and in the economy at large, all the developed countries in the world are attempting to transform education in the sciences.

In spite of this effort to reform science education, very little progress has been made in achieving a significant, meaningful, and effective change in the teaching of science (Hurd, 2000). What is lacking are policies governing the decisions that have to be made about the role that science and technology play in human affairs; as well as policies that control the integration of the personal-social and social-civic aspects of science as we know it today (p. 282). Journals publishing research done in science education rarely include policy studies (p. 282). It seems as though science education researchers value statistical and quantitative policy studies more than those which are qualitative and philosophical in nature. Consequently, the science reform movements of today have stalled and have become politicized (p. 283).

Furthermore, science and technology have become such an integral part of the economic, political, and social fibre of life today that the practice of science has to undergo a paradigm shift if it is to be effective (Hurd, 2000). Such a paradigm shift would demand "creative forms of collaboration between scientists and society" (p. 283), and the involvement of a wider range of science disciplines and competencies in this process. The relationship between science and technology in this 21st century is perceived as "two sides of a single coin" (p. 283). In the past, science teaching has paid close attention to the inquiry processes implemented by scientists. There was then the expectation that students would emulate and practice these processes and thereby learn to "think like a scientist" (p. 283). It is now evident, however, that there is no single, standard method for the practice of science. In the 21st century, "research in the sciences is viewed more as a craft or an art, more like problem solving" (p. 283).

The educational challenge of this Information Age is "to access, synthesize, codify and interpret science information into a working knowledge that can be used in personal and civic contexts" (Hurd, 2000, p. 284). In order to become relevant in the lives of students, science courses have to include problems, projects, investigations, and experiments in applied settings that engage students in their own affairs and in their own cultures: science as a *lived* curriculum. Under such conditions, a teacher takes on the role of a "coach, guide, consultant, mentor or co-learner" rather than functioning as a "talking head" (p. 286).

During the past 40 years, the only policy practised by all countries attempting to transform science education has been "learning to learn" (Hurd, 2000, p. 284). This adage is an endeavour to introduce students to scientific knowledge in such a manner that they are enabled to traverse life on their own, maintaining themselves always as responsible citizens and productive workers. Hopefully a commitment to this adage will instil in students a vision of lifelong learning, which will equip them with the human capital needed to meet the challenges of career changes at all stages in their lives. In today's world, where there is an oversupply of information, the scarce resource is the cognitive ability to make sense of this information (Gilbert, 2005). Science curriculum reform for this Information Age, therefore, should provide citizens with "higher order thinking skills, intellectual adaptability and the ability to manage knowledge" (Hurd, 2000, p. 285). According to Osborne (2013), "the failure to transform science education for the needs of the 21st century is a consequence of a lack of a good model of scientific reasoning, and of a body of expertise about how to assess such higher order cognitive competencies" (p. 265). The overarching goal of science education today should be "to educate scientists who will be at home in society and to educate a society that will be at home with Science" (National Science Foundation, 1970).

Some science educators are of the opinion that the contexts we use for teaching science are inappropriate and irrelevant (Hurd, 2000). However, there is little evidence to suggest that "more relevant" contexts lead to a better understanding of ideas, although they may lead to greater motivation (Osborne, 2009). There seems to be an underlying problem here that has not been addressed adequately: the teaching of science needs to be rethought and reconsidered, especially at the first-year undergraduate level (Alberts, 2005). There appears to be a flaw at the very core of what we are doing as science educators (2005), and since the first point of contact our students have with science is through the science textbooks they use in their classrooms, we should take a closer look at them.

Science Textbooks.

Science textbooks, used widely in science classrooms throughout Canada, play a major role in science instruction. They have a tremendous influence on how students perceive the scientific enterprise (DiGisi & Willett, 1995; Scruggs, 1988; Yager, 1983). In many science classrooms, the majority of the instructional support, beyond the teacher, is provided by the prescribed textbooks; consequently, much of the scientific information that students receive comes from science textbooks (Mayer, 1983). Some researchers believe that in the majority of classrooms these textbooks serve as the most authentic source of authoritative scientific information and in many cases “actually becomes the curriculum” (Chiappetta, Sethna, & Fillman, 1993, p. 787; Stake & Easley, 1978). Science textbooks often become the sole resource in a science class because their vocabulary is specific to the grade level of the students and because they present information in a concise and succinct manner (Lentz & Evans, 2005). Research studies (for example, Chiappetta, Fillman, & Sethna, 1991), have shown that textbooks fail to attend to all aspects of scientific literacy because they tend to focus more on the content of science (substantive knowledge) rather than the process of science (procedural knowledge). By addressing the appropriate procedural knowledge whenever possible, teachers could have a moderating impact on the didactic power of textbooks. Another alternative would be to add APL, which focuses on both substantive and procedural knowledge. Substantive knowledge (based on facts), and procedural knowledge (based on the thinking behind the doing of science), are the two essential requirements of problem solving in science education.

Science in Higher Education

There are many important issues affecting science education in higher education all over the world. In the United States, the President’s Council of Advisors on Science and Technology

(2011), sounded the warning that due to the waning number of students choosing careers in science, technology, engineering, and mathematics (STEM), this country is losing its competitive edge within the market place (Stuart, 2000). “More and more American students are pursuing nonscience majors, creating a shortage of eligible candidates for science-related careers” (Kumar, 2003, p. 125). A survey of the literature highlights the following as transitions in science education occurring in higher education since the year 2000 especially in the United States:

> A shift in the research paradigm; the rising tide of commercialism; more stringent human subject research regulations; teaching for a scientific workforce; the instructional technology invasion; pressure to participate in science teacher preparation; ethnic and gender inequalities; a proliferation of research disciplines; and fading public influence.(p. 125)

The key trends listed above will be examined in more detail below.

There is a paradigm shift in science research from hypothesis-based to more problem-based and interdisciplinary research which addresses specific societal needs, such as improving the health of humans, protecting the environment, sustaining natural resources or developing response readiness to bioterrorism (Kumar, 2003). Until now, most of the science policies developed and implemented in the United States centred on university research. With the rising tide of commercialism, the general public today is less inclined “to invest in isolated research at universities, which often ends up in obscure refereed journals gathering dust on library shelves” (p. 126). There is, therefore, a change in the emphasis from specific university-based research to more research projects involving university-industry partnerships, which are expected to create new jobs and businesses. In other words, policymakers are conveying a clear message to universities that they have to start commercializing their discoveries so as to enhance economic development locally (Schmidt, 2002).

In Canada, the greatest development is science education since 2006, has been the effort to decolonize the Pan-Canadian Science Framework by viewing Indigenous knowledge as foundational to the understanding of the physical world (Aikenhead, 2006). This work encouraged Indigenous communities and leaders to negotiate appropriate modifications to the science curricula. There is now evidence (Aikenhead & Elliott, 2010) that such integrated science courses have been implemented and found to be both feasible and educationally sound.

It is critically important for science educators, to prepare their students for the work environment of today which is dominated by science and technology (Kumar & Chubin, 2000; National Research Council, 1995). A public campaign was initiated in Canada in order to prepare students for the modern scientific workforce (Birchard, 2001, A6). Canadian universities put out an advertisement, aimed at widening the pool of the workforce choosing science-and-technology- related careers, which read, “Learn how to think critically and you’ll be prepared for almost anything - and land a good job besides” (Birchard, 2001, A6). This advertisement was aimed at “encouraging college students to learn science for the sake of education and not just for preparation for professional careers in medicine and engineering” (Kumar, 2003, p. 128).

By using information technology, it is possible to develop technology-based instructional resources (Flynn, A. B., Caron, J., Laroche, J., Daviau-Duguay, M., Marcoux, C., & Richard, G., 2014): the impetus to do so is an encroaching challenge on the status quo of college science teachers. One way to confront this challenge is for these teachers to use technology to increase the quality of their instruction (Flynn, A. B., 2012; Flynn, A. B., & Amellal, D. G., 2016). University science professors are sometimes pressured, and always encouraged, to work with local public schools and thereby put in time and effort to enhance teacher education (Kumar, 2003).

Just as the current rates of advancement in the fields of science and technology are alarming, the retention of students in these fields at the higher education level is likewise an issue of international concern. Many groups of science researchers and individual scientists have been concerned lately about the decline in the number of students pursuing the sciences and the resulting dwindling numbers of science graduates (Scott, 2005; Sjoberg & Schreiner, 2005). Several researchers have attempted to explore and fathom the reasons for the choices that students make, at crucial stages in schools, concerning science and science careers (Frost, Reiss, & Frost, 2005; Gilborn & Gipps, 1996; Reiss, 1998; and Rodrigues, Jindal-Snape, & Snape, 2011). Consequently, several reasons have been offered to explain why students are choosing to opt out of science and the science career pathways. In certain parts of the world, such as Thailand, it is with a sense of urgency that education policy is enforcing that "science education must aim to enhance students' capability and interest in Science" (Office of National Education Commission, 2003, cited in Yuenyong & Narjaikaew, 2009, p. 341).

Studies (Drew, 2011) have found that, approximately 40% of students who initially choose science majors switch to other fields. In this 21st century there is a dire need – greater than ever before – for highly educated people in science who are equipped to meet the economic, environmental and technological challenges of this age (Alberts, 2005; Bordt et al., 2001;). Yet, the reality that science educators face today is a decline in the number of students choosing to pursue the study of science (Kind, Jones, & Barmby, 2007). As early as 1989, Green summarized his concerns about this situation by categorically stating:

> Not only do the sciences have the highest defection rates of any undergraduate major, they also have the lowest rates of recruitment from any other major. In short, science departments lose a huge proportion of their potential clients – the academically-able and intellectually-motivated students who enter college with a genuine interest in studying science. (p. 478)

This shows that retention in science is a problem at the undergraduate level. What can be contributing to this problem of retaining students in science?

I would like to explore this further by looking locally, at the retention picture at the University of Alberta in Science, Education (Science and Math), and Engineering? In 2005/2006, 15.4% of the student intake in these faculties did not return to the University of Alberta in 2006/2007 (R. Wimmer, personal communications, December 3, 2008). The vast majority of students in the Faculties of Engineering and Education continued into the second year of their programs (e.g., in engineering 80% of males and 72% of females). In the Faculty of Science, however, the numbers dropped to 69% of males and 66% of females, displaying a gender discrepancy.

In 2013-2014, although women constituted 47.3% of the total labour force in Canada, they made up only12.5% of the workforce in general physical science professions, and only 20.5% in general engineering professions (Canadian Association of University Teachers, 2013-2014). It is not surprising, then, that the underrepresentation of women in math, science, and engineering fields has been the focus of much research in the last two decades. Enrolment in accredited engineering undergraduate programs reached 73,035 in 2013, an increase of 24.1% since 2009 (Engineers Canada, 2014). A total of 13,363 undergraduate degrees in engineering were awarded in 2013 (Engineers Canada, 2014). This was an increase of 23.9% from 2009. Although the share of women awarded undergraduate degrees increased from 17.6% in 2009 to 18.3% in 2013, these women account for only one-fifth of the total population who completed a bachelor's degree in engineering in 2013 (Engineers Canada, 2014). A large volume of research has been published in the last decade in Canada, exploring why there was such a low proportion of women, especially at the decision-making levels, in the sciences. Myths are still rampant as to what influences the choice of career, as well as what guides the progress women make in science (Madill, et al., 2007). It is evident that factors that affect retention vary between faculties. It is also probable that faculty and gender have far-reaching, important, and interconnected, ramifications for retention. There may be several factors contributing to the lack of retention in science, and specifically the factors inhibiting retention in chemistry - the science discipline which is the focus of my study.

Area of Focus

One of the most important of these factors is the problem of knowledge transfer: that is, there is a distinct gap between what is done at the frontiers of cutting-edge chemical research and what gets taught as "chemistry" at the high school and university undergraduate levels. Since there is dynamic research occurring in the field of chemistry, the gap between what researchers are discovering and what is being taught at the first-year undergraduate level in universities and colleges is widening. Scientists, especially chemists, tend to publish their most recent research work in academic science journals, which are referred to as *primary literature*. In primary literature, the scientists who did the research are the ones who write about it. There are ways in which primary literature can be used to bridge the gap between public knowledge and knowledge at the frontiers of scientific inquiry. Although courses based on primary literature, or that use primary literature, are quite common at the university senior undergraduate level (Epstein, 1970; Houde, 2000; Levine, 2001; Smith, 2001;), many students at the freshmen level have limited experience reading primary literature, despite the fact that the ability to read and evaluate this form of information is crucial for their success in graduate school or in the scientific work force (Janick-Buckner, 1997; Kuldell, 2003; Peck, 2004;). Using primary literature may be a good way of nurturing a "capacity for scientific ways of thinking among students" (Baram-Tsabari &

Yarden, 2005, p. 403). Primary literature can, therefore, have a positive role in teaching subject matter in chemistry, as well as in developing scientific literacy.

To be scientifically literate, students must have the reading ability to evaluate the print-based information presented to them, as well as the writing ability to communicate their thoughts to others (Holliday, Yore, & Alvermann, 1994). In high schools, students obtain scientific information from textbooks, newspapers, or some other form of media, and review articles in popular journals (Wellington, 1991). Scientists are not usually the authors of these texts that are therefore classified as *secondary literature*. The usual method, by which scientists communicate their results to peers in their field, is through peer-reviewed journal articles called *primary literature*. When science instructors introduced primary literature into their classes, they noticed that their students found them hard to understand (Muench, 2000). However, primary literature can be adapted to simpler forms suitable for various target populations. Recently, texts described as "adapted primary literature" (APL) have been introduced by science instructors into their classes.

What is Adapted Primary Literature (APL)?

APL, based on the research of Baram-Tsabari and Yarden (2005), is an adaptation of primary literature that purposefully retains its canonical form and authentic results. However, for this research project, the APL was written to be understood by first-year undergraduate students. The rationale underlying such adaptations is that they are more consistent with the nature of scientific inquiry than are undergraduate chemistry or biology textbooks (Smith, 2001; Muench, 2000). APL usually starts with questions about phenomena rather than answers to be learned and reproduced (Kuldell, 2003). University undergraduate students who learn through APL may develop a capacity for scientific ways of thinking and reasoning (Baram-Tsabari & Yarden, 2005).

Incorporating primary scientific literature into classes is not easy (Muench, 2000) but its use has much value since this literature has a unique potential to instruct students on the nature of scientific reasoning and communication. APL, which is used in my study, has the capacity to promote scientific literacy and enhance scientific ways of thinking (Yarden et al., 2001) because of its:

> (a) acquaintance with the rationale for research plans; (b) exposure to research methods and their suitability to research questions; (c) acquaintance with the language and structure of scientific communication; (d) potential to develop the ability to critically assert the goals and conclusions of scientific research; (e) exposure to problems in certain disciplines; and (f) acquaintance with the continuation of the scientific research process. (p. 190)

Baram-Tsabari and Yarden (2005), examining how text genre affects the formation of scientific literacy, found that students reading APL showed greater understanding of the methods of scientific inquiry and demonstrated more critical thinking and application abilities than those who read only secondary literature. However, students who read secondary literature understood the text better and were more positive about what they read. Employing APL alongside science textbooks may provide a good combination.

Evidence

What is Evidence?

Evidence is difficult to define. Philosophers have differing views about evidence. In a dictionary of philosophy, Mautner (1999, p. 184) defined evidence as “that which provides a ground for a belief or a theory.” In another philosophical dictionary, Audi (1999, p. 293), described evidence as “information bearing on the truth or falsity of a proposition.” A proposition is a statement, usually in the form of a declarative sentence which contains a “that” phrase either explicitly or implicitly (Miller & Fredericks, 2003). Karl Popper (1959), in his book *The Logic of Scientific Discovery*, asserted that it is evidence that upholds the truth in a proposition: this is particularly true when the proposition in question is based on an experimental observation or experience. Evidence, a concept central to the empirical sciences, is used when referring to data or observations that are put forth to support or refute a scientific hypothesis. Whether to believe in a scientific hypothesis or theory depends upon the quantity and character of the evidence in its favour. The idea of "concepts of evidence" was introduced by Duggan and Gott in 1995.

Concepts of Evidence (CoEs). The study of science requires the acquisition both of substantive knowledge and skills. Substantive knowledge includes the understanding of facts, concepts, and theories of science. Concepts of Evidence refer to concepts associated with procedural understanding. Skills may include knowing how to use a thermometer, how to draw a graph, how to set up a distillation apparatus, or how to focus a microscope. It is not always recognized that skills have a distinct knowledge base that is connected with the understanding of scientific evidence. Skills must be exercised within an understanding of ideas such as variables and their manipulation, accuracy, fair testing, and the validity and reliability of data. It is these ideas that Roberts and Gott (2004) and Gott and Duggan (1995) termed "Concepts of Evidence," and which I have called CoE (Concept of Evidence) for short. According to Roberts and Gott (2004), Concepts of Evidence supply "the underpinning ideas about how evidence can be *collected*, *verified*, *analyzed,* and *interpreted*" (p. 11). Concepts of Evidence are concepts underlying the doing of science in relation to the evidence as a whole. They form a knowledge base akin to the substantive knowledge that has traditionally been seen as the heart of science. These Concepts of Evidence were also termed "conceptions of scientific evidence" (Taylor & Dana, 2003). Questions about how APL can be employed to test students' understanding of

Concepts of Evidence form the basis of this research project.

Purpose of the Study

In this study, I have used APL as a tool to determine whether the initial, pre-instructional understanding that students have of the Concepts of Evidence – concepts which govern the collection, handling, verification, analysis and representation of data – can be affected by pedagogical intervention. The purpose of this study, therefore, is to find out if pedagogical interventions can affect the understanding that first-year undergraduate students have of these Concepts of Evidence. The teaching intervention employed in this study was a video, which was sent electronically to each participant, when the individual had completed the Pre-Test. The video was 32 minutes long and depicted a professor in the chemistry department at the University of Alberta, who was the principal investigator of the research project described in the APL. In the video, this professor described a few of the experiments as well some of the Concepts of Evidence used implicitly in this work.

Research Questions

At its inception, this research project was initiated to explore the question: Can Concepts of Evidence be taught or learned; and if so, what are some of the factors that influence their teaching and their learning*?* However, as the study progressed, four research questions emerged:

1. What difference did the teaching intervention make to the initial understanding that students held of these Concepts of Evidence (i.e. procedural knowledge) and of the substantive knowledge covered in the APL?
2. a) What were the experiences of students as they read the APL, and went through all the stages of this research activity?

   b) What were the experiences of the instructors as they read the APL?

3. What are the views of first-year undergraduate students and their instructors about introducing Concepts of Evidence as part of their curriculum in first-year chemistry courses?
4. What is the relationship, if any, between students' understanding of the Concepts of Evidence and pedagogy?

Researcher Background

I am a person of East Indian descent born to East Indian parents who were both high school science teachers. Science and the teaching of science were discussed and highly valued in my family. At an early age, I was struck by the way my parents as teachers were making a difference in the lives of their students. My parents' sustained dedication, care, and concern were greatly respected and valued by their students. I thought teaching was such a worthwhile profession because it could change, enlighten, and improve the lives of others. This has left a deep impression on me and is the main reason I have chosen to be a science educator, and why I have persisted on this long, personal journey.

Although I was born in India, I grew up in Sri Lanka, Bahrain and various parts of Africa, ranging from Ethiopia in the north to South Africa in the southernmost tip. I did my first degree in India, majoring in chemistry, and did a Master of Science in biochemistry at the University of London, England. I had some very valuable and cherished experiences pursuing research in my biochemistry programs that left a lasting impression on me of the value of empirical research. Working in the biochemistry labs highlighted in my mind the importance of collecting and handling good data. It impressed on me the skills required to evaluate data before they can be turned into evidence supporting the findings in a research project.

When my family and I immigrated to Canada, I did a Bachelor of Education in order to be able to teach in this country. I taught for about six years in Edmonton, and then started an Masters in Education degree. While doing my MEd, I undertook a study that addressed the question of how high school chemistry textbooks treat the nature of scientific evidence, where I conducted an examination of the three most widely used high school chemistry textbooks in Canada for their inclusion of Concepts of Evidence. I found that the treatment of Concepts of Evidence varied widely across textbooks and, within textbooks, across topics. I argued that there needs to be a comprehensive and systematic curriculum for procedural knowledge much as there exists one for substantive knowledge. Teachers may need guidance in the recognition and teaching of these Concepts of Evidence, which are the building blocks of procedural knowledge.

Structure of Dissertation

This dissertation is divided into five chapters. The first chapter is an introduction to the background of the study and explores its necessity. The second chapter is a detailed literature review of work done on evidence in general and Concepts of Evidence in particular. This chapter will explore the role of APL in bridging the gap between science in the forefront of research and what gets taught as science at the undergraduate level in postsecondary institutions. In Chapter Three, the design of the research project will be discussed with emphasis on the qualitative methodology and the methods employed. The results and findings, discussed in Chapter Four, include a qualitative analysis of the evidence survey and the interview data. The evidence survey, comprised of a pre-test and a post-test, were completed by all the first-year undergraduate student participants. The interview data, on the other hand, were obtained from interviewing these student participants as well as their instructors. Chapter Five will be devoted to addressing the research questions. The conclusions and implications for teachers, for research and policy, and the further research questions that arise will be reflected on this last chapter.

CHAPTER TWO

LITERATURE REVIEW

Prelude: The Transplanting

When the rice seedlings (known locally as *njaru*) reach a height of 10-15cms, they are ready to be transplanted. The rice seedlings are plucked and tied up in small bundles. Transplantation day arrives, and women do most of the work associated with this process. The older women came out dressed in *mundu* and *chatta*, while the younger girls donned long skirts and blouses. They were all wearing a head-cloth or a small umbrella on their heads in order to protect them from the scorching heat of the sun. The women walked down the *varambu* and each got to one segment. Once in the segment each woman picked up a bundle of rice seedlings. All the women working in the same segment lined up in single file and spaced themselves so that they covered the length from one end of the segment to the opposite end. Each woman then opened up the bundle and separated out the seedlings which were transplanted into one particular segment of the paddy field. Bent double, up to their ankles in mud, each woman made a hole in the soil with a finger, planted the seedling into the hole and covered the delicate roots with the soil. As the women toiled they joined in to sing. They sang songs that had familiar tunes but they created their own verses from occurrences in their daily lives. By the end of the day, the women had transplanted the clumped rice seedlings into the paddy fields in neat straight lines. Then it was time to go home.

The above vignette is analogous, in part, to the place given to literature review in research. Often knowledge related to a research project occurs in “bundles” analogous to the bundles of rice seedlings. When the researcher has identified various “bundles” of knowledge from reviewing the literature, the individual “seedlings” representing sections of knowledge have to be neatly organized, and given its appropriate place sequentially. This is what is accomplished in a literature review

Introduction

I began this literature review by building on the references I obtained during research for my Master's thesis, where I examined how evidence was portrayed in chemistry textbooks used in high schools across Canada. For my dissertation, I had to look for literature on students' understanding of evidence. I was fortunate to attend academic conferences and participate in some panel presentations where various aspects of evidence – the philosophical, theatrical, scientific, and medical – were discussed and debated. These experiences broadened my understanding of the applications and usage of evidence. Here I have explored the background and theoretical basis for this research under seven headings: Trends and Themes in Science; What is Knowledge?; The Coordination of Theory and Evidence; A Model for Science; The Portrayal of Science in the Literature; Factors Influencing How First-Year Undergraduate Students Learn; and, Science as a Social Practice. I conclude the chapter with a summary of the interrelations and connections among these seven bodies of knowledge.

Trends and Themes in Science

In most countries throughout the world, science is given an important place in the secondary school curriculum. Science has attained such a taken-for-granted status that it no longer requires any justification for its inclusion in curriculum writings (Millar & Driver, 1987). If one looks critically at the history of school science education, however, one can observe a cycle of periods in which the method or process of science has been strongly emphasized, alternating with periods in which content featured more prominently. One could envision this as pendulum swings when sometimes, like in the early 1900s, substantive knowledge (with a stress on content in learning science) has been emphasized, while at other times, such as the 1950s and 1960s, procedural knowledge (emphasis on process in learning science) has been dominant. This pattern of a recurrent cycle is observed very generally, across national boundaries (1987).

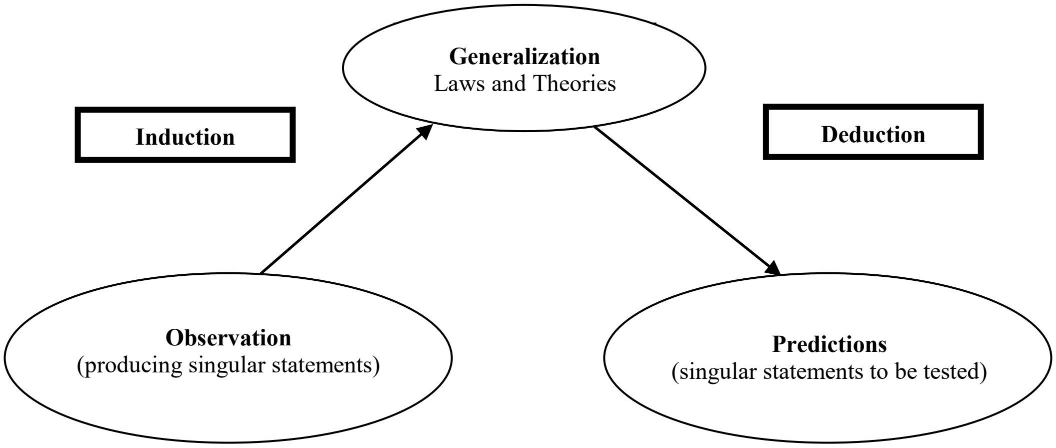
The “Scientific Method”

According to process science, the methods of science are considered to be a sequence or hierarchy of processes, beginning with something simple like observation and leading to more complex classification, inference, and finally ending in a hypothesis. Here, observation takes precedence over theory, and this is basically a very simple, unsophisticated, inductive view of science (Millar, 1991). According to the inductive method, scientists first make many observations of nature, with the objective of making a few powerful statements about how nature works. These powerful generalisations are called laws or theories. The main problem with induction is that we can never be certain that inductive generalisations will hold or be valid. We may have proof that several inductive generalisations are valid, only to find that the next one we examine goes against that generalisation. Philosophers have attempted, in vain, to generate a “logic of induction” – that is, a set of rules that regulate the path from a gathering of separate, specific instances to a generalisation that is applicable universally (Millar 1991). Such philosophical attempts have not been successful and have, to date, ended in failure (Chalmers, 1982). In the inductive method, the observation of what happens in nature is the authority. If an idea clashes with what happens in nature, then that idea must be changed or discarded.

The traditional view of the “scientific method” has been summarized in Figure 1 (Hodson, 1998, p. 12). It depicts how observations are made as statements. These singular statements are turned into universal statements or generalisations by induction. Laws and theories, once established, make some predictions by deduction. These predictions are also singular statements which can then be tested.

Figure 1

The Traditional View of the Scientific Method



Adapted from (Hodson, 1998, p. 12)

The main criticism of the inductive view of science is the great asymmetry that exists between proof and disproof. For example, no matter how many instances have clearly demonstrated that a particular generalisation is true, one single non-proof or disconfirming observation can falsify that generalisation. This led Popper (1959) to propose a hypothetico-deductive view of Science, where he argues that the method of science entails a "rigorous testing of hypotheses, in an attempt to falsify them" (Millar, 1991, p. 46). According to Popper (1959), it is only by falsifying a theory that exists today, that we can progress towards a better one tomorrow. This is referred to in the philosophy of science as Popper's falsification, and has serious ramifications.

A crucial problem arises here, because most philosophers of science (including Popper), maintain that all observations are theory-laden (Hanson, 1958). If this is true, then what we observe depends on the conceptual apparatus, and the prior theories that we bring with us to the task of observation. One has to presume then, that even a disconfirming observation is not exactly “pure,” but depends on the theories held by the observer. There is a chance, therefore, that the theories held by the observer are the ones which are erroneous, and not the theory being tested. Consequently, the idea that a single observation can conclusively falsify a well-established theory is a bit far-fetched (Mulkay, 1979).

To sum up, we can conclude that to date there is no agreement among historians, philosophers, and sociologists of science about whether or not science has a method. If science does have a method, these experts are unable to articulate what it is. However, within this pool of uncertainty there are substantial points of agreement that need to be noted, and this is how the literature supports them (Millar, 1994). For example, most researchers in this area agree that “scientific enquiry cannot be portrayed as rule following but involves the exercise of skills” (p. 167). In the process of observation itself, one has to decide what to observe, what observations to pay attention to, what interpretations and inferences to draw, what conclusions to draw from empirical data, and even how many times the experiment has to be replicated (Collins, 1985). Such decisions cannot be pinned down to a set of rules, nor can they be attributed, on retrospect, to an intuitive set of rules in the head. As Millar pointed out, much of what scientists know about “the method (or methods) they follow is tacit; doing science is like practising a craft” (Millar, 1994, p. 167; Polanyi, 1958; Ravetz, 1971). This has serious implications for the way science, and more specifically scientific enquiry, is taught.

Towards Attainable Goals.

What are some attainable goals for eradicating distorting usages of “process?” First, there is a need to distinguish between specific skills which can be improved by pedagogy, and the more general “processes” which cannot. The “processes” of science are “features of general cognition” which are not specifically linked to science alone and ... “we all use them routinely without instruction" (Millar, 1989, p. 56). The term "process" when used in Science, refers to activities such as observing, classifying, inferring, hypothesizing, and so on, while the word "skill" is used for more specific activities such as "reading a thermometer to within 1^0C" (p. 48). Skills can be taught and improved, and therefore, they can be assessed. Second, we need to be clear that what we are talking about under observation is "scientific observing" and not observing in general. Third, we need a clearer understanding of the stages in developing these science skills. Thoughtful and detailed analyses such as the one done by Norris (1984, 1985), on observation, are needed to replace the superficial descriptions of "processes" which are in the literature, if we are to make any headway in separating the achievable from the illusive.

What is Knowledge?

Knowledge is comprised of facts, information, description, and skills acquired through experience or education. It includes both the theoretical and practical understanding of a subject. In addition, knowledge is the awareness or familiarity gained by the experience of a fact or a situation. In philosophy, the study of knowledge is referred to as *epistemology,* a word that comes from two Greek words, *episteme* which means knowledge, and *logos* meaning study of. The study of knowledge is something philosophers have undertaken, "for as long as philosophy has been around" (Pardi, 2011). Epistemologists study various aspects of knowledge such as what constitutes knowledge, what kinds of things can we know, what are the limits to what we can know, and even whether it is possible to actually know anything at all (Pardi, 2011). Although no single agreed-upon definition of knowledge exists, Plato defined it as "justified true belief" (Fine, 2003, p. 5). This means there are three criteria to be met if a statement is to be knowledge: it must be justified, true, and believed.

The knowledge explosion of the 21st century is unprecedented, and has far-reaching effects on the knowledge economy. Numerous attempts have been made in the past to describe knowledge systematically (de Jong & Ferguson-Hessler, 1996). Some of these trials have been based on cognitive theories, whereas others have been developed in order to function as a foundation for instructional design theories. Still others have been formulated from an epistemological standpoint. In the epistemological approaches, the elements of the knowledge base are classified by the function they perform in the fulfillment of a target task; in other words, they are task-dependent. This means that different types of ontologies (classifications) of knowledge types are devised for different types of tasks (Gott, 1989). In addition, this also implies that the same elements of subject matter may be characterized by different ontologies, based on the task they perform within one particular domain (de Jong, de Hoog, & Schreiber, 1988) – an implication that has some interesting consequences when one attempts to classify knowledge.

Types of Knowledge

One such attempt was made by de Jong and Ferguson-Hessler (1996), whose research was in the field of physics, where problem-solving and experimental work have a crucial role. From their work on problem-solving, these authors were able to differentiate four different types of knowledge:

Conceptual/Declarative/Substantive Knowledge.

This is knowledge about facts, concepts, and principles, which are applicable within a particular domain. Within a problem-solving domain, conceptual knowledge acts as additional information that solvers can add to the problem, and which they can readily use to get the solution. Roberts and Gott (2000) defined “substantive ideas” as “the concepts which form the underpinning structure of the subjects," for example, ideas such as energy, force, photosynthesis, or solubility (p. 83). To sum up, substantive knowledge in science is the understanding of ideas based on facts, laws, and principles.

Procedural Knowledge.

Procedural knowledge is the understanding needed to put science into practice. "*It is the thinking behind the doing,*" which includes the understanding of a set of ideas which are complementary to substantive understanding, but which are related to the "knowing how" of science (Gott & Duggan, 1995, p. 26). "It encompasses the actions or manipulations that are valid within a domain" (de Jong & Ferguson-Hessler, 1996, p. 106). Lubben and Millar (1996) took the stance that procedural knowledge in science is "knowing how to carry out practical tasks" (p. 957). For example, in a plant growth study, procedural knowledge does not refer to the measuring itself, but to the decisions that have to be made about what to measure, how often, and over what period of time. In empirical work, basic experimental skills are needed to succeed: but they are not enough on their own. "There is a higher cognitive dimension, which has been called procedural understanding, necessary for the correct application of those skills" (Roberts & Gott, 1999, p. 22). The content of procedural knowledge is not documented properly, although it is a kind of understanding "in its own right*,"* equivalent to substantive understanding (Gott & Duggan, 1995, p. 26). The knowledge base of procedural knowledge is not taught systematically or with rigour. To address this absence, Gott and his co-researchers made a tentative list of 80 or so ideas, which they have collectively called the *Concepts of Evidence* (Gott, Duggan, & Roberts, 2003). These Concepts of Evidence have been classified into 21 categories and include ideas about the uncertainty of data (Buffler, Allie, & Lubben, 2001), ideas needed "for understanding measurement and data processing, presentation and analysis which may be considered to be part of the mathematics curriculum but which are essential for understanding evidence" (Roberts et al., 2010, p. 379). It is important, therefore, to have a clearly laid out knowledge base for procedural knowledge, and to incorporate this into the K-12 science curriculum, so that it is addressed systematically and progressively during the teaching of school science.

When knowledge is properly classified, the subject matter concerned will be organized and regulated so effectively that it will direct the process of learning in that field. When the field of study is science, the acquisition of knowledge becomes even more complex, because in this field, although knowledge advances are substantial theoretically, it can be argued that some great strides forward are attained empirically, because science is a practical subject.

Acquiring Knowledge in Science

It is important, therefore, to make students aware of the nature of these empirical engagements with science, and to prepare them to carry out their own scientific investigations. This is a desirable goal promoted by the science curricula in various countries of the world. Yet, scientific knowledge, celebrated as one of the greatest achievements of western culture, would seem, if anything, to be increasingly under attack (Appleyard, 1992). Science educators, as Osborne (1996b) pointed out, have the responsibility "to interpret and communicate a complex but fascinating subject and to provide constructive learning opportunities for children" (p. 271). This is an onerous task, and one that science educators need to meet on a daily basis. As Hanson (1958) maintained, what the child or individual brings to the practical work determines what they get out of it. For example, what students bring to the process of observation determines what they are able to see. Therefore, if students are to benefit from practical work, teachers have to ascertain that these students spend much more time developing the appropriate conceptual framework that is required, so that they can be equipped to accurately interpret the data (Hodson, 1990). "Pupils need to spend more time interacting with ideas and less time interacting with apparatus" (p. 34). Moreira (1980), for example, in a study of practical physics lessons conducted in a number of British Secondary Schools, found that students very often perform experiments in class, equipped with only a very basic idea of what they are doing. They often have no notion of the objectives of the experiment, or of the underlying theoretical concepts, or the reasons why a certain procedure is adopted in a particular experiment. They seem to be carrying out these experiments in a cook-book fashion, like following recipes (Moreira, 1980). Such activities are not only a waste of time, but are confusing and counter-productive.

In science it is not sufficient to be concerned only with acquiring knowledge, but one has to ensure that this knowledge is valid. This certainty is not generated by faith or by intuition, but by knowledge that could be described as well as communicated. Plato (Joyce, 2004) had three tests for certain knowledge: (1) Test One: You must *Believe* the Statement; (2) Test Two: Your Belief has to be *True*; (3) Test Three: Your True Belief must be *Justified*. Knowledge that passes all three tests above is referred to as evidence. Knowledge, therefore, requires justification, and justification is a matter of having sufficient evidence (2004).

The Coordination of Theory and Evidence

Kuhn, Amsel, and O'Loughlin (1988) focused on the development in young people of the capacity to engage in scientific reasoning. They examined the ability of students at different ages to evaluate given (or self-proposed) theories using given evidence, and considered the influence that the students' own theories had on this process. Kuhn et al. introduced students to a problem, elicited their theories about it, and then asked them to say whether the given pieces of data supported or conflicted with a particular "theory" and to explain their reasoning. The authors concluded that the coordination of theory and evidence is subject to developmental change and that many young students do not consider the possibility that their theories might be false or that alternative theories might exist. The ability to separate data and explanation (or “theory”) takes time to develop.

This requirement for time was emphasized further in D. Kuhn’s review article (1989) that compared scientific thinking with everyday thinking. The view that Kuhn endorsed in this article is that the coordination of theories and evidence is at the epicentre of scientific thinking. “A central premise underlying science is that scientific theories stand in relation to actual or potential bodies of evidence against which they can be evaluated” (Kuhn, Amsel, & O’Loughlin, 1988, p. 3). According to D. Kuhn, scientists are: (a) able to express in their own words a theory they believe in; (b) aware of what evidence does and could support their theory, and what evidence would be contradictory to it; and (c) able to give a justification for why the coordination of available theories and evidence has led them to accept a particular theory and reject others claiming to account for the same phenomena. These skills in coordinating theories and evidence are the most valuable, necessary, and general skills that define scientific thinking.

The Coordination of Theory and Evidence is Progressive

Several investigators, including Deanna Kuhn and her coworkers (Klahr & Dunbar, 1988; Schauble & Kuhn, 1989), exploring the processes of coordination of theory and evidence, have indicated that these processes are significantly different in the child, the lay adult and the scientist. If these processes are different, it would be useful to have a developmental framework, in order to conceptualize these variations.

Initially, Kuhn et al. (1988) looked at the processes by which theories are revised when they meet with new evidence. The theories of the subjects were initially assessed, and they were then asked to generate evidence and also evaluate various forms of evidence that they were provided. The evidence generated or provided, were sometimes in agreement with, and at other times in conflict with, the subjects' own theories. As anticipated, the theoretical beliefs of the subjects influenced the generation and evaluation of evidence. This influence acted in distinctly different ways in the child, the lay adult, and the scientist. There were two ways displayed by the subjects to maintain the alignment between theory and evidence: (1) the adjustment of evidence to fit a theory; and (2) the adjustment of theory to reduce its inconsistency with evidence. Children often accepted a hypothesis as true "based on minimal evidence and overlooked conflicting evidence" (Kuhn, 1989, p. 682). Although the children generated as much data as the adults, their experiments were less well designed. The children were also less able to make generalizations or universal theories from their data; they tended to be content with local interpretations, i.e., inferences consistent with the last result generated (p. 682). The professional scientist displays full coordination of theories and evidence, and the interaction of theory and evidence is elevated to the level of conscious control. It is in the above-listed important respects that the professional scientist differs from the child scientist.

Later work by Karmiloff-Smith (1984, 1986, 1988) illustrated that for a number of developmental phenomena, the coordination of theory and data is progressive. Second, the progression that she described is an exchange between data-bound and theory-bound procedures. Third, Karmiloff-Smith stressed that the coordination of theory and evidence is not something that is "mastered once and for all in some domain-free manner" (Kuhn, 1989, p. 682). As shown by her data, mastery can be achieved in some simple tasks such as narration, but in more complex tasks, even adults may have attained only incomplete mastery. However, thinking in scientific contexts, as described above, represents only a minute portion of human thinking.

In further research, Kuhn (1991) employed a broader range of contexts, in which the skills that have been described above were utilized. Her seminal research question was: “In what ways might the weaknesses in reasoning that have been described manifest themselves in everyday thought?” (Kuhn, 1989, p. 683) Her participants were in four age groups – teens, 20s, 40s, and 60s. Kuhn asked her participants “to describe their own causal theories and to relate evidence to them” (p. 683). The results showed that theory and evidence were fused together into a narrative script of “how it happens” (p. 683). The participants sometimes appeared to be unable to produce alternate theories, or to provide counterevidence for a particular theory. In coordinating theory and evidence, the key elements to be recognized by subjects are: (1) the possibility of alternate theories; and (2) the possibility of evidence that does not fit a theory.
In this research project no age or sex differences were significant, but education had a sustained, positive effect. College students showed a marked superiority to non-college subjects. One has to urge caution in this kind of interpretation, because “education level is to some degree an effect as well as a cause of reasoning ability” (Kuhn, 1989, p. 684). One automatically wonders what skills the subjects need in order to effectively coordinate theories and evidence.

Theory in science leads to laws and principles which are the cornerstones of what is referred to as substantive knowledge. Data collected empirically, on the other hand, are evaluated using Concepts of Evidence which are the building blocks of what is known as procedural knowledge. If the data meet the rigors of the evaluation they are turned into evidence: if they do not, they return to being just data. Such ideas cumulated into a simple model for science.

A Model for Science

Based on the premise that “practical work should be an integral part of the science curriculum which mirrors, reinforces and augments the rest of the course,” Gott and Mashiter

(1991) proposed a basic, simplified model for science (Gott & Duggan, 1995, p. 25). In Gott and Mashiter's model, conceptual understanding represents understanding the ideas in this subject which are based on facts, laws and principles. Conceptual understanding is sometimes referred to as "substantive knowledge" or "declarative knowledge" (p. 26). Procedural understanding, on the other hand, is the comprehension of a set of ideas which is "complementary to conceptual understanding," but which appertains to the "knowing how" (p. 26) of science, required to put science into practice. According to Gott and Duggan,*"It is the thinking behind the doing"* (p. 26). For example, in taking the measurements in a study on plant growth, procedural understanding refers to "the decisions that have to be made about what to measure, how often and over what" period of time (p. 26). The content of procedural knowledge is not well documented, especially in the school science curriculum. In spite of the fact that procedural knowledge *can* be a means of learning or learning about a concept," it is in addition, a type of understanding "*in its own right*" (p. 26).

Gott and Mashiter (1991) defined a problem as "a task for which the pupil cannot immediately see an answer or recall a routine method for finding it" (p. 58). In order to solve a problem or investigate, there are "a set of procedures" which must both be understood and implemented appropriately (p. 58). These necessary procedures or skills may be summarized as (Gott & Mashiter, 1991):

1) Identifying the important variables;
2) Deciding on their status – independent, dependent or control;
3) Controlling variables;
4) Deciding on the scale of quantities used;
5) Choosing the range and number of measurements, their accuracy and reliability;
6) Selecting appropriate tabulation and display. (p. 58)

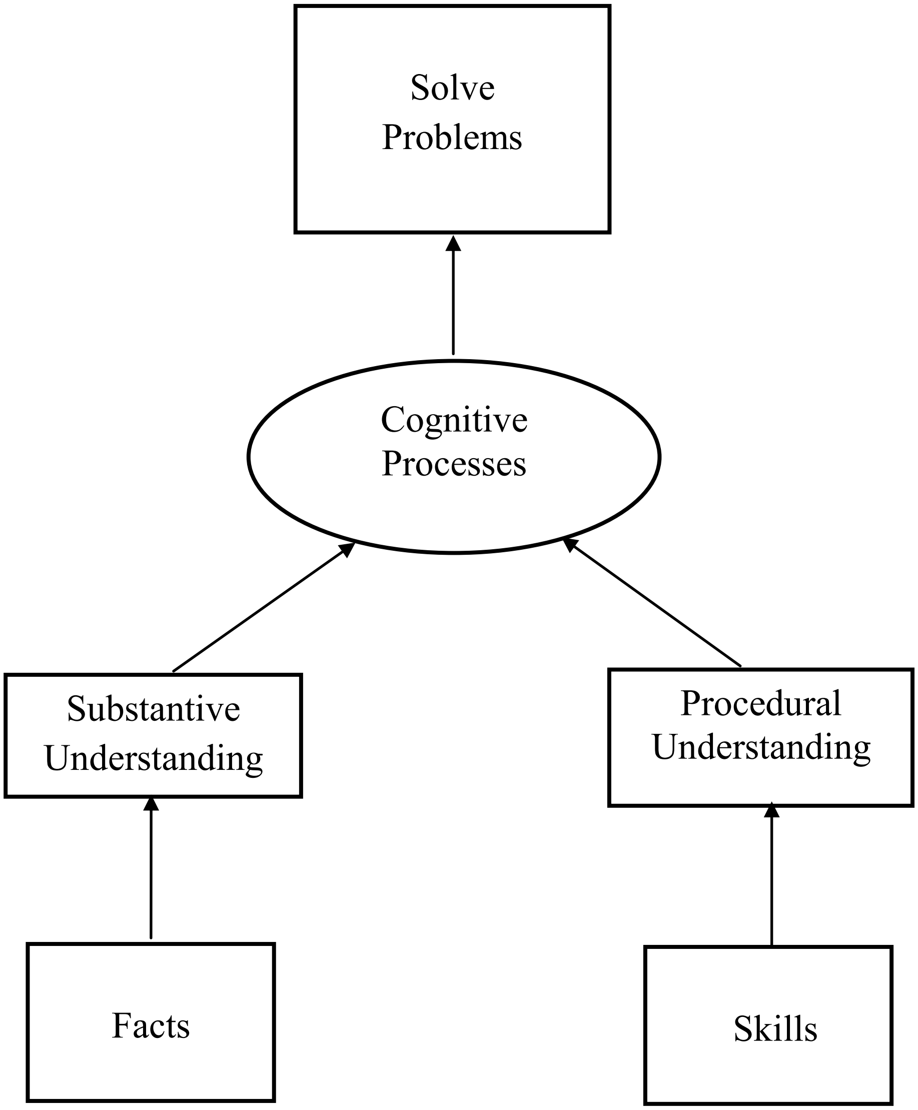
Procedural knowledge in science may be described as "the understanding of how to put all these specific skills together" via the identification and operationalization of variables and the display and interpretation of data" (p. 59).

According to this model (see Figure 2), the cognitive processes that are used to solve problems entail an interaction of conceptual and procedural understanding. The model does not imply that these two types of understanding are "mutually exclusive" (p. 25), on the contrary, emphasizes how they are inextricably linked: neither is complete without the other. Conceptual and procedural understandings are not totally independent of each other: a certain level of substantive knowledge is needed to perform "most procedural aspects of Science," and conversely, some procedural knowledge is required to put substantive understanding into practice (Gott & Duggan, 1995, p. 27). In addition, the substantive understanding and the procedural understanding are depicted in Figure 2 as unidirectional and both feed into the cognitive processes. These cognitive processes involve dynamic interactions between the selection and application of facts, skills, substantive knowledge, and procedural knowledge (1995). The ability to obtain and process the information needed to tackle a problem successfully rests in these cognitive processes. Gott and Mashiter (1991) have not specified what exactly these cognitive processes are and therefore, this requires further investigation. It is noteworthy however, that factors such as motivation, context, or student's expectation, which can have a profound effect on performance, have not been accounted for, in this model.

Gott and Duggan (1995) coined the phrase *Concepts of Evidence* to signify the concepts that are the building blocks of, and are associated with, procedural knowledge. These Concepts of Evidence were developed based on the four major stages of investigations: design of the task; measurement; data handling; and evaluation of the complete task (1995). The term Concepts of Evidence accentuates the importance of procedural understanding and "the concepts underlying the doing of science in relation to the evidence as a whole" (p. 30). It would be timely, therefore,

Figure 2

A MODEL FOR SCIENCE



(Based on Gott & Mashiter, 1991)

to clarify this notion of Concepts of Evidence by using examples encountered by first-year undergraduate students taking science or chemistry courses.

Examples of the dynamic interaction between Substantive and Procedural Knowledge

Some examples of Concepts of Evidence encountered by students learning first-year undergraduate science will help to clarify how the basic model of science put forward by Gott and Duggan (1995) can be applied.

For Students doing First-Year Undergraduate Chemistry Courses.

A very good example of how the model of science proposed by Gott is applied is provided by some first year undergraduate chemistry laboratory exercises utilized in courses such as Chemistry 101: Introductory University Chemistry I. Suppose that the problem was to find out whether the mass of copper deposited on zinc electrodes was related to the concentrations of the copper sulphate solutions in which they were submerged.

To carry out this investigation, zinc strips acting as the electrode, are placed in copper sulphate solutions which are blue in colour. The five zinc electrodes are of the same dimension and have to be weighed initially. In order to do this a balance has to be chosen (*procedural knowledge*, CoE#7 – The choice of an instrument for measuring a datum) (See Appendix G, p. 344), which measures mass accurately and precisely to the first decimal place. Before commencing to weigh the zinc strips, the balance must be calibrated and checked for zero error (*procedural knowledge*, CoE#5 – Instruments: Calibration and error). If there is a positive or a negative zero error a corresponding zero error correction must be applied to each reading. The balance must also be checked to ensure that the underlying relationship between increase in mass weighed and the reading displayed on the balance is a linear one (*procedural knowledge*, CoE#4 – Instruments: Underlying relationships). This will also help to establish the reliability and validity of each measurement of mass using the balance (*procedural knowledge*, CoE#10 – Reliability and validity of a datum). All the zinc plates weighed 50g each (*procedural knowledge*, Design: Choosing values).

The students need to know some basic electrochemistry (*substantive knowledge*), the branch of chemistry that deals with reactions in which electrons are transferred known as oxidation-reduction reactions or redox reactions. All metals and their ions (i.e., charged metal entity) have their own characteristic electrode potentials. The standard electrode potential (*substantive knowledge*) is an intrinsic measure of the tendency of a chemical species (in this case metal ions) to acquire electrons and thereby be reduced. It is measured in volts (V) or millivolts (mV). The more positive the standard electrode potential, the greater is the tendency of the chemical species to be reduced (i.e., the stronger it is as an oxidizing agent) (*substantive knowledge*). Conversely, the more negative the standard electrode of a chemical species, the lesser is its tendency to be reduced (i.e., the stronger it is as a reducing agent). Based on their differing standard electrode potentials, metals and their corresponding ions, can be arranged in a series of half-reactions(an electrode immersed in a solution containing ions of the same metal) ranging from the strongest oxidizing agent (SOAs) to the strongest reducing agents (SRAs). Such an arrangement is provided to first year undergraduate students, as the "Table of Selected Standard Electrode Potentials" in their data booklets (*substantive knowledge*).

From the data collected, as the concentration of copper sulphate solution increases, the mass of copper deposited on the zinc plate also exhibits a similar increase. Therefore, in response to the research question asked at the beginning, one must reiterate that the masses of copper deposited on the zinc electrodes were definitely related to the concentrations of the copper sulphate solutions in which they were submerged. The concentrations of the copper sulphate solutions, therefore, have an effect on the masses of copper deposited on the zinc electrodes. There is a preponderance of Concepts of Evidence (*procedural knowledge*), which have a role in ensuring the validity, reliability, and reproducibility of the investigation undertaken here. If one is to undertake such an investigation or interpret the results from such an experiment, these Concepts of Evidence come to the fore, having a substantial role in their design, implementation, and mode of interpretation.

Exploring Concepts of Evidence as a Framework for Building Procedural Knowledge

The science curriculum in high school and first-year undergraduate level do not specify which Concepts of Evidence (procedural knowledge) can be taught at various points as the substantive knowledge is covered. It would be very necessary to explore the various concepts of knowledge (e.g., Matter as solutions, acids, bases and gases; Electrode potentials and redox half-reactions) included in the respective curricula and see which of them, and at what points, would lend themselves to be used as a framework for introducing certain Concepts of Evidence. In this way the instructors can build on their students' procedural knowledge as well, while they cover the curricula containing mainly substantive knowledge.

As seen in the undergraduate example described earlier, there are some Concepts of Evidence associated with certain applications, as well as with practical and experimental tasks, at various points in the substantive knowledge covered in the curriculum. These points could be signalled out in the curriculum as a framework for introducing and developing procedural knowledge into the teaching of science. At these points in the curriculum, instructors could specifically teach and address these Concepts of Evidence. This framework could be used to address the procedural knowledge progressively, beginning with the simpler ones and leading to the more complex ones, because some Concepts of Evidence are more difficult than others. Such a rudimentary effort to include Concepts of Evidence in the teaching of science would be greatly enhanced if the procedural knowledge and the theory behind it would be included, spelt out and laid down specifically - just as substantive knowledge is laid out - in the curriculum document.

Implications for Pedagogy and the Teaching of Science

There has been a call, in the last two decades, to change the way we teach Science from the more teacher-centered delivery to a more student-engaging method that fundamentally recognizes the way people learn (National Research Council, 2000). The human mind can only assimilate at a certain controlled rate and comprehend only a limited amount of new information (Eberlein, Kampmeier, Minderhout, Moog, Platt, Varma-Nelson and White, 2008). Transferring knowledge more or less directly from teacher to student by using the common "teaching by telling" method has been found to be ineffective for many students (Eberlein et al., 2008, p. 263). In science courses, as the cognitive load increases, it becomes increasingly important that students are involved actively in their own learning. Students must construct for themselves a "workable understanding of sophisticated concepts" (p. 263), and they must be encouraged to develop their own higher-order thinking skills. The main active learning pedagogies employed in science education are: team problem-based learning; process-oriented guided inquiry learning; and peer-led team learning.

Each of the three active learning pedagogies mentioned above is based on the tenets of *social constructivism*. In science education, constructivism maintains that "knowledge is constructed in the mind of the learner by the learner" (Eberlein, Kampmeier, Minderhout, Moog, Platt, Varma-Nelson and White, 2008, p. 262; Bodner, 1986). According to social constructivism, this building of knowledge is brought about by cooperative social interactions. Social constructivism also implies that teaching by only delivering lectures, fails to exploit fully the unique characteristic of the educational milieu today, in which students are prepared to make a mutual effort to master both course content and the pedagogy required to accomplish this most effectively.

The substantive knowledge, when dealing with the topic of electrochemistry, for example, tends to be very abstract and difficult to visualize for students. However, combining this conceptual knowledge with an investigation, or experiment or demonstration geared towards answering a specific research question, draws in several concepts of procedural knowledge that are observable and applicable. This eradicates the remoteness of the substantive knowledge and makes it very relevant with pliable consequences which can be manipulated experimentally using inquiry-based learning. The abstract *substantive knowledge* in the topic of electrochemistry is therefore, made alive and observable through this investigation and the introduction of a host of very relevant and suitable Concepts of Evidence (*procedural knowledge*).

Since most teachers in higher education, have not experienced any of the pedagogies which promote active learning, there is a significant barrier for them to accept any of these unfamiliar methodologies. Consequently, there is a gap between what is now known about the process of learning through science education research, and the way that science is often taught at present (Anderson, 2007). Although some faculty may recognize the merit of active-learning methods, there is an activation energy required to try something new, and a fear of the consequences of failure that must be overcome to plunge into the unknown. This is a source of major concern for educators interested in reforming science education by adopting appropriate active learning strategies. Yet in science education, it has been shown that “faculty who embrace the need for instructional change and conquer the initial steep learning curve, find the experience transformative" (Eberlein, Kampmeier, Minderhout, Moog, Platt, Varma-Nelson and White, 2008, p. 270).

Erroneous Conceptions or Misconceptions

The research done by cognitive psychologists and science educators, shows that children and adults "hold a variety of naive, intuitive conceptions – usually misconceptions – about how the world works" (Kuhn, 1989, p. 675; Carey, 1985, 1986; Chi & Roscoe, 2002; Nicoll, 2001). These conceptions, although wrong, have been shown not only to be powerful, but they conspicuously resist instruction (Viennot, 1979; Costu & Ayas, 2005). Science educators, therefore, have to find out what these wrong conceptions are, and work towards modifying them (Davis, 2001), rather than just superimposing the new correct concepts on the misconceptions. It is difficult to make contact with these wrong concepts that students hold, because these mental models contain some explanation of how "key elements in the model relate to one another and how phenomena operate" (Kuhn, 1989, p. 675).

Conceptual Change

Leading on from this view of scientific thinking as mental models is the idea that the development of scientific thinking can be envisaged as a progression of partially correct theories within different conceptual domains. Some researchers such as Glaser (1984) and (Chi, 2000b), have suggested that such a progression is the central feature of cognitive development, and others (Carey, 1985; Chi, 2000a; Krupa, Selman & Jacquette, 1985; Vosniadou, 2007) have actually identified such patterns of development for certain scientific concepts. Parallels have also been drawn (Carey, 1986; Gruber, 1973; Limon, 2002) between the progression of scientific understanding in students and the progression of scientific understanding in the history of science. The child's development of understanding has been compared to T. Kuhn's (1962) account of the history of science as a succession of paradigms of unequal duration. The term "child" (Carey, 1986; Limon, 2002) is used here in a generic sense, and is not qualified with factors that could have an impact such as gender, geography, race, and socio-economic status: herein lies a limitation of this work. Just as a paradigm is preserved in spite of conflicting evidence, so too there is weak restructuring (known as assimilation) in a child's conceptual change (Carey, 1986; Limon, 2002). However, a strong restructuring (known as accommodation) in a child's conceptual change (Carey, 1986; Geelan, 2000; Limon, 2002) is equivalent to a paradigm shift as framed by T. Kuhn, and is the more important kind of restructuring that occurs in conceptual change.

Such restructuring, in the strong sense, involves changes in the core concepts of the theory, and the interrelation of these concepts (Kuhn, 1989). Because of the strong restructuring that occurs, it may not be possible to directly translate the core concepts in the new theory into those of the old theory. Both the child and the scientist, according to Kuhn, get a deeper understanding of the world through constructing and revising, "a succession of models or paradigms that replace one another" (p. 675). This is sometimes referred to as the child-as-scientist metaphor. However, what exactly is the *process* by which these mental models or theories or paradigms are revised? If the revision of theories is the focal point of cognitive development, should we not know more of how it occurs?

The initial, landmark work in the accommodation of a scientific concept was done by Posner, Strike, Hewson and Gertzog (1982). According to them, "learning is a rational activity... best viewed as a process of conceptual change" (p. 212). They also proposed the Conceptual Change Model, derived from the philosophy of science (p. 211), which stipulates that four conditions must be satisfied if a conceptual change is to take place in the understanding that an individual has about a construct. These four conditions will be explored in Chapter Five, and will be illustrated with an illustrative example from the students' interviews.

Researchers in the area of conceptual change (Chi, 2000a; Vosniadou, 2007) have little to say about the processes by which theories are initially constructed and then revised as a way of learning about the world. Carey (1985) has claimed that these progressions of understanding have to be identified and described before any investigation of their mechanism can commence. She, therefore, focussed her work on complex domains which are very rich in content. In these complex domains, characterizing the conceptual content of successive theories requires detailed description and obscures any question of mechanism.

Since a mechanism for conceptual change has not been forthcoming, theorists have offered four competing views of what constitutes this process. According to Vosniadou (2007), prior knowledge is very important to learning. He envisions conceptual change as a process in which students are able to synthesize models in their minds based on their existing explanatory frameworks. He argued that conceptual change is a gradual process resulting in a progression of mental models.

Chi and Roscoe (2002) on the other hand, perceived conceptual change as a repairing of one's misconceptions. Sometimes students have naive but faulty conceptions which must be identified and then rectified. From Chi and Roscoe's perspective, misconceptions are just concepts which have been miscategorized. Therefore, conceptual change is achieved by the reassignment of concepts to the correct categories.

According to diSessa (2002), conceptual change is attained when diverse, fragmented naive knowledge is organized cognitively into complex systems in the minds of students. Ivarsson, Schoultz and Saljo (2002) took the stance that conceptual change is the appropriation of suitable intellectual tools. When students change the ways in which they use these tools in differing contexts, conceptual change occurs at the societal level.

Although there are four differing views on the process of conceptual change as discussed above, progress has been made in suggesting the implications for pedagogy. Wiser and Amin (2002) promoted the use of computer models coupled with verbal interactions where the teacher enables the scaffolding of ideas. Niaz, Aguilera, Maza and Liendo (2002) suggested that if students are offered the opportunity to argue and discuss their ideas then their "understanding can go beyond the simple regurgitation of experimental detail" (p. 523). Another recommendation is that science educators should pay more attention to the history and philosophy of science during instruction.

In my study, I have used scientific literature as a tool, to test first-year undergraduate students' understanding of these Concepts of Evidence. Consequently, it is vital to examine the various ways in which science has been portrayed in the literature, and to select the most suitable form for my research project.

The Portrayal of Science in the Literature

As I moved closer to identifying my central focus in this inquiry, it became clearer to me what I needed to learn to enter the conversation in an informed manner and develop it in a worthwhile direction. I did an extensive review of the literature, available in English, on primary literature and how it was being used in classrooms at high school and university undergraduate levels. I found that primary literature was used in order to enhance scientific literacy and bridge the gap that exists between public knowledge and scientific knowledge at the frontiers of scientific inquiry. I considered how secondary literature compares with primary literature in accomplishing the above-mentioned goals. I looked at how review articles and newspaper science play out as sources of secondary literature in promoting scientific literacy.

Primary Literature

Scientists, by convention, publish their research work in academic journals to make their findings known to academia and the world at large; these publications are classified as primary literature (Yarden, Brill, & Falk, 2001). Professional journals are the venue scientists use to discuss their results and findings and to describe advancements in various fields made as a result of their work. These contributions and advancements can be described adequately only by using scientific reasoning. “Reading and analyzing primary literature is an authentic scientific cognitive activity, as scientists’ conclusions are grounded in the theoretical and empirical work of other scientists” (Chinn & Malhotra, 2002; Dunbar, 1995; Falk, Brill, & Yarden, 2008, p. 1842). Primary literature, therefore, reveals the progress of scientific reasoning, and thereby promotes scientific literacy. Another advantage of introducing primary literature into the science classroom may be that students will acquire “a realistic understanding of the nature of science, by exposing them to the authentic inquiry activity of reading scientific texts which has been reported to have many benefits” (Epstein, 1970; Falk, Brill, & Yarden, 2008, p. 1842; Janick-Buckner, 1997; Muench, 2000;). The scientists who carried out the research are themselves the authors of primary literature. Very often in primary literature, scientists use the jargon and terminology associated with each field. At both university and college levels, courses based on reading primary literature, or courses accompanied by the usage of primary literature, are a common occurrence, and have been well documented. The advantages and disadvantages of using primary literature, therefore, are explored next.

The Merits and Demerits of Using Primary Literature

Primary literature is of prime importance in maintaining the purposefully argumentative structure of science (Pall, 2000). Being subject to peer-review, primary literature plays a crucial role in preserving the integrity of the scientific enterprise; that is, in “keeping the scientific enterprise honest” (Abelson, 1980, p. 61). One has to read and understand primary literature to fully fathom the process of science and to appreciate how different it is from other disciplines, and from other human accomplishments (Pall, 2000). Although undergraduate science majors may be exposed to scientific literature, it is seldom that non-science majors get a taste of it. Consequently, non-science majors and citizens at large remain unaware of the “deliberative structure of Science” (p. 256).

Pall (2000) addressed this shortfall in a general education science course for non-science majors by introducing students to primary literature in human genetics. The general education human genetics course required students to write a term paper in which they compared “an article in the peer-reviewed scientific literature with a popular article on the same subject” (p. 256). Pall reported that the quality of the term papers was quite high, demonstrating a considerable amount of thought on the part of the students. He did notice, however, that his students had to be warned that when the scientific data and opinion are in conflict, the data should always be given priority.

The greatest value of primary literature is that it familiarises students and instructs them on “the nature of scientific reasoning and communication” (Muench, 2000, p. 255). Primary literature can enhance the learning of science in the classroom in that it can act as a tool for teaching scientific literacy. Brill and Yarden (2003) suggest that “learning through research papers may be one way to provide a stimulus for question-asking by high school students and results in higher thinking levels and uniqueness” (p. 266). This said, it is not easy to incorporate scientific primary literature into the classes (Muench, 2000), and choosing the appropriate primary literature article can be a challenge.

Choosing Primary Literature

One of the key factors that influence the success of introducing primary literature is the actual choice of the peer-reviewed journal article (Muench, 2000). Nucleic acid damage induced by ultraviolet (UV) light had been studied extensively in DNA (Wenger & Loppnow, 2004). However, the photochemistry of RNA has been much less intensively researched, despite the fact that uracil (a nucleobase found only in RNA) appears to be the most photochemically active nucleobase (Wenger & Loppnow, 2004).

Muench (2000) described a number of criteria to take into consideration when selecting papers to use in undergraduate classes, such as: (1) What are the educational goals of the course?; (2) What are the goals of the assignments?; (3) What concepts must the reader understand?; (4) Is the relationship between the data and the conclusion simple and direct or complex and abstract?; (5) Fringe benefits: What else can students learn from this paper? (p. 256–259). All the above questions were addressed and carefully thought out when choosing the primary literature paper to adapt for the first-year undergraduate students in my study.

Student Responses to the Introduction of Primary Literature in Advanced Courses

Janick-Buckner (1997) offered an advanced level cell biology course designed to incorporate “the critical reading and examination of Primary Literature” (p. 29). Students’ responses to the course were evaluated toward the end of the course using open-ended questions and a standardized evaluation form. On the whole, the students agreed that their critical reading, writing, and analytical skills improved as a result of what they experienced in the course. Many students were of the opinion that this course was a great help to their undergraduate research. Not only were they equipped to design their own experiments, they were better interpreters of their own data. An extremely important positive outcome of the course was that students gained confidence "in their ability to reason, research, and apply knowledge" (p. 32). According to these authors, the fact that they were able to understand state-of-the-art research in molecular and cellular biology, and were well-informed to recognize the limitations to certain methods, made many of the students very self-assured and satisfied with their accomplishments.

Secondary Literature

In most cases, the text genres that high school students deal with are found in textbooks, review articles, and popular research news obtained from the media (Yarden, Brill, & Falk, 2001). These texts contain "descriptions of scientific research performed by scientists who are usually not the authors." Therefore, the texts are classified as secondary literature. Popular reports of science are the primary means to provide science education for non-scientists (Korpan et al., 1997). Consequently, the ability to understand and interpret such reports serves as a good indicator—a litmus test—of scientific literacy, which is a prime goal of science education (Phillips & Norris, 1999). Learning science through newspapers is regarded as an example of informal learning, yet the relationship it bears to formal learning is hard to fathom (Wellington, 1991). It is a debatable point whether the science present in newspapers "can be of value in formal science education if used carefully and critically" (p. 363).

Adapted Primary Literature

A research article in which "the basic typical structure of the original article was retained, along with the use of passive voice," is defined as *Adapted Primary Literature* (APL) (Baram-Tsabari & Yarden, 2005, p. 405). APL is an educational genre which was first designed specifically to teach biology at the high school level (Falk, Brill, & Yarden, 2008). The basic canonical structure of the original primary literature article is preserved, as are the original results, figures, and illustrations. Since the alterations implemented were included to make the text simpler, but not to modify it significantly, this altered version of the research article is referred to as APL (p. 405). A distinct gap exists between the “highly professional nature of research articles and the cognitive level of high school students” (Yarden, Brill, & Falk, 2001, p. 191). In the first example of APL that I created therefore, the original journal article underwent “a certain degree of processing” (p. 191) aimed at simplifying it so that it would be suitable for an audience of first-year undergraduate students.

Characteristics of APL

APL is a unique genre created to enhance the effectiveness of primary scientific literature (PSL) in the learning of science, especially at the high school level (Yarden, 2009). The new genre intends to present “real science” in high schools, “and to promote important aspects of high school students’ scientific literacy that are harder to achieve using textbooks or JRV [Journalistic Reported Version]” (Yarden, 2009, p. 309). The authors of APL are science educators and scientists, and the target audience is usually students doing high school science (Yarden et al., 2001). APL uses mainly the argumentative genre, as opposed to the narrative or expository genres usually found in science textbooks (Norris et al., 2009b). Since APL maintains the canonical structure of primary scientific literature (Baram-Tsabari & Yarden, 2005; Yarden et al., 2001), there is enough evidence in its content to support conclusions that are laid down specifically (Yarden et al., 2001; Falk & Yarden, 2009). In APL, science is presented as being uncertain and subject to change, and alterations are expected when new advances are made (Falk & Yarden, 2009; Yarden et al., 2001). It has long been established that “many science curricula … do not take into account the practical reasoning required in scientific knowledge production” (Norris, 1992, p. 196). It is, therefore, plausible to assume that using the argumentative genre, prevalent in APL, may ensure that scientific reasoning is incorporated into the school curriculum (Yarden, 2009).

APL has many promising characteristics (Ford, 2009):

> It has the ability to introduce complex, authentic, and cutting edge examples of actual scientific arguments to high school learners. It expands the written text genres available to learners, and assists them in understanding the many different ways that scientific ideas can be conveyed. (p. 389)

According to Ford (2009), APL can enhance not only the learning of science content, but the epistemology of science as well.

However, some science educators are not convinced that APL is the only way to introduce scientific reasoning into the classroom (Osborne, 2009). Some are of the opinion that since APL is secondary literature, its relevance and accuracy depends on the understanding and the writing abilities of the person constructing the APL. The structure of the APL is tied to how the writer understands and interprets what is presented in a journal article as primary literature. In this regard, an APL can have the same weaknesses as a newspaper article (Osborne, 2009; Norris et al., 2009a). This comparison between APL and a newspaper article is addressed later in this chapter and in Chapter Three.

Some science educators are of the opinion that instructors, especially at the first-year undergraduate level, should adhere to using primary literature (Hoskins et al., 2007; Levine, 2001; Smith, 2001), in their attempts to engage their students as active participants in the learning process. Instructors could choose articles that are relevant, interesting and written in a language and style suitable for first-year undergraduate students, in spite of the fact that this student group may find the task of reading the primary literature challenging (Smith, 2001). Besides, primary literature is readily available because that is the way scientists communicate with their peers in describing their research and findings (Yarden, Norris & Phillips, 2015). Furthermore, science instructors might not have the time or expertise necessary to write good APLs. It could be argued that this task could be assigned to science educators doing research in Science Education. Besides, a good APL depends on choosing good primary literature, which many science instructors feel students should be exposed to in the first place in order to increase their critical thinking skills.

From a pedagogical perspective, APL seems like a sound approach for teaching science because it uses scaffolding, which ensures that the learning is anchored on past experience of the learner. Such anchoring is a fundamental aspect of the constructivist theory of learning (Driver, Asoko, Leach, Mortimer & Scott, 1994). Scaffolding, therefore, allows children to accomplish what they could not have achieved on their own: it would have taken too much time and effort. Computer technology is an excellent means of anchoring learning to meaningful experiences. It must be noted that the use of multimedia will have an impact on the development of the human brain, especially when such techniques are used with children and adolescents, whose brains are still developing and maturing (Forrester & Jantzie, n.d.).

The meanings of words provided in the right-hand column of the APL, would act as a forum for the students to check their initial, pre-instructional understanding of these scientific terms. They could make a mental check and address any misconceptions they may hold of these terms. Because their meanings are provided, students are not required to commit these terms to rote memorization.

Symposium and Special Issue on APL

A special issue of the journal *Research in Science Education* focussed on the reading of scientific texts, directing its attention particularly to APL. This special issue developed from a symposium that took place at the 2008 annual meeting of the National Association of Research in Science Teaching (NARST) (Phillips et al., 2008). Researchers examining the use of a variety of settings and formats of primary scientific literature for learning science at the high school level gathered at the symposium. The special issue includes nine papers presented at the symposium and the discussions that followed the presentations.

Reading as Inquiry. The Phillips and Norris (2009) study, which was the first paper in this special issue, views the languages of science and school science to be extensively different. They trace this difference to the simplistic view that “reading is being able to say the words correctly,” a theory that is prevalent in science education and most of schooling (p. 317). They argued:

> Reading and writing are inextricably linked to the very nature and fabric of science and, by extension, to learning science. Take them away and there goes science and proper science learning also, just as surely as removing observation, measurement, and experiment would destroy science and proper science learning. (p. 226)

A study by Suppe (1998) found that journal articles from various disciplines of science had a common structure as far as the organization of the papers was concerned. All the papers showed certain “speech acts” such as: presenting the data obtained from observations; relating how the observations made were pertinent to some scientific problem under investigation; elaborating on the methods used for data collection and analysis; giving and justifying a certain interpretation of the data collected; and specifying, admitting, and challenging alternative interpretations (Phillips & Norris, 2009, p. 315). The “speech acts” described above “create an argumentative structure” that is found in all the data-based papers examined by Suppe (Phillips & Norris, 2009, p. 315). A study by Norris (1992) suggested that one of the professed goals of science education should be to teach students “the shape that the arguments need to take to support given conclusions” in science, what he called the "‘justificatory shape’ within science” (p. 316). The language of science, therefore, is inextricably associated with the argumentation genre.

Students with a simplistic view believe that “reading is being able to say the words correctly” (Phillips & Norris, 2009, p. 317). Background knowledge does not aid those with a simplistic view of reading; a science background benefits only those students who have developed “sound reading strategies” (p. 317). In another study, Norris and Phillips (2008) reasoned that the best way to think of reading is as “an inquiry process,” where readers draw meaning from text “by integrating relevant text information with their relevant background knowledge” (p. 318). This type of integration leads to the proper interpretation of the text. Background knowledge plays an important role in the process of interpretation because it is responsible for making “inferential links” between pre-existing knowledge and the text (p. 318). This puts reading in the limelight as a constructive process (Phillips & Norris, 2009). As such, reading entails many of the same mental activities that constitute the core of what science is, and what doing science involves. If science educators do not stress the importance of reading science text as inquiry, they are likely to promote the simple view of reading, thereby devaluing both the “complexity” as well as the “importance” of reading in science (p. 318). How then can we draw students away from the simplistic view of reading?

According to Phillips & Norris (2009), one way to achieve this would be to implement reading as inquiry as a part of the school science curriculum by using APL in the classroom. Since APL preserves the canonical structure of scientific papers, but is written to be understood by high school students and undergraduate freshmen, it resembles the language of science and can be easily understood by these students. The chasm that exists between the language of science and the language of school science can be successfully bridged by using APL in the classroom.

Promises and Challenges for the use of APL in Science Curricula. APL has great potential as a tool for engaging students with science as well as enabling them to learn about the reasoning as well as the stylistic practices of science. Three themes from the special issue of the journal *Research in Science Education* (NARST, 2008) were discussed in Ford (2009):

(1) Authenticity - was defined as “an adherence to scientific reasoning and practice within science education” (p. 385). So far, “authentic” experiences for science learners have been conveyed using school science curricula that are activity-based and that also put emphasis on empirical approaches (Ford, 2009). Critical inquiry into scientific arguments that have been written down, however, is still given prime focus in the practice of science (Hand et al., 2003; Norris & Phillips, 2003). In Ford (2009), it is emphasised that:

> Scientists use written texts as a means to convey their finalized arguments, and the ability to evaluate these written arguments is both a sign of expertise in science and also a central component of understanding the nature of science. To understand the ideas of science, one must necessarily understand the conventions of language used to present and support these ideas to an informed audience. Understanding the written language of science is arguably an authentic practice, whether within a scientific community or within a school science community. (p. 386)

Authenticity not only provides a purpose for the activity-based aspects of school science, it also justifies the understanding of science as a set of balanced practices that include “reasoning and argument,” “the development of explanations,” and above all, “the sharing of ideas with colleagues” (Ford, 2009, p. 387). Using APL in the classroom would allow students to realize that scientists engage in reading and writing. Moreover, from reading APL, students would recognize that scientists can be “quite passionate about Science” and attempt to assure others of the genuineness of their ideas using “carefully written arguments” (p. 387). From reading APL, students would also recognize that texts are written in science not just to convey substantive knowledge/facts, but even more frequently to develop scientific arguments and to expose these arguments to the evaluation of other scientists, especially those working in the same field. As students become more familiar with the usage of APL, they will begin to see how scientific arguments have certain similarities in their structure (Phillips & Norris, 2009), and how they can apply these arguments.

(2) Implementation of APL—The selection of papers to be covered as APL needs careful consideration of matters such as engaging students and motivating them, as well as seizing "authentic scientific practices" for the success of the curricula (p. 388). It is important to implement discipline-specific science reading strategies and epistemologies. Teachers also need to be supported in their attempts to implement APL, and to do this successfully "we need to develop a research-based pedagogy of science reading inquiry" (p. 388).

(3) Extending the concept of APL to older and younger populations—The APL research that is emerging has been done mainly in high school classrooms. It would be interesting to surmise what experiences similar to APL would be possible with younger students or students at a college level. One can imagine endless possibilities if APL is extended into college classrooms, especially in non-major and introductory science courses, as well as courses intended for pre-service science teachers (Ford, 2009, p. 389). These APL enactments would be precursors of the real thing—primary literature—which could be interspersed in the advanced science courses. In my study, I prepared and tested an APL at the first-year university undergraduate level.

The Potential of Adapted Primary Literature (APL) for Learning. Osborne (2009) welcomes this collection of papers (in a special issue of the journal *Research in Science Education* developed from the symposium held at the annual meeting of the National Association of Research in Science Teaching in 2008) which challenge "what it means to learn science," stating that all these papers convey the theme that using APL in high school adds value to the learning of science in these classrooms (p. 397). Osborne raises three issues that question the underlying assumptions in Yarden et al. (2009). The first issue is whether APL is the most appropriate way to introduce a layperson to science. Osborne explains that during the early 19th century, when contemporary science was presented in the salons of London, the presentations were highly popular because the educated person of that day could easily understand them. Because of its cumulative nature, contemporary science is remote to the understanding of the layperson. Adapting the primary literature to decrease the background knowledge the layperson will require to build a suitable understanding is one option, but is it the most appropriate approach? One could argue that the text students are most frequently exposed to, are reports of science in the media, and that training students to read and understand this genre should be part of their science education.

The second issue identified in Osborne (2009) is "In what sense can reading be perceived as an act of inquiry?" Reading as inquiry was first promoted in Norris and Phillips (2008). Reading as inquiry was based on the premise that reading is seen as a constructive process in which meaning must be deciphered from text "by forging inferential links between the reader's background knowledge and the text" (Osborne, 2009, p. 399). One of the shortcomings of science teaching in schools today is an undue emphasis on experimental tasks involving the "manipulation of the material world," instead of surveying with students the "interpretation of data" (Watson et al., 2004) or the "language of science" (Lemke, 1990). Inquiry, to put it briefly, "is a process requiring a multiplicity of actions" (Osborne, 2009, p. 400). In Norris and Phillips (2003) it is stated that both reading and writing are activities vital to the process of inquiry. However, as argued in Osborne (2009): "… how many of these activities are necessary or how many are sufficient to constitute engaging in inquiry is open to debate" (p. 400). Although reading, according to Norris and Phillips, is certainly "a process of inquiring into meaning," and interpreting text may "require the reader to generate inferences," it is hard to fathom how reading itself "can be seen *as inquiry*" (Osborne, 2009, p. 400). Osborne illustrates this point by constructing an argument. According to him, reading is a central process of scientific inquiry—in other words, "it is part of the process of inquiry, but it is not inquiry itself" (p. 400).

The third issue mentioned in Osborne (2009) is that most of the authors in the special issue of *Research in Science Education* (NARST, 2008) vouch for providing students with opportunities to participate in "authentic scientific practices" (p. 400). An argument that runs through most of the papers in this collection suggests that students, by "engaging with authentic or adapted authentic texts," will gain a better understanding of both the content knowledge in science and the process of science (p. 400). Osborne cautions that there are two dangers here.

There is a prevalent assumption that the most efficient/effective way of learning science is by "replicating the practices of scientists themselves" (Osborne, 2009, p. 400). Osborne argued that just as a good sports commentator does not have to be good at sports himself, so too one does not have to be a scientist to get a good understanding of the major, basic theories that science has to offer about the world around us.

Another point made by Osborne is that authenticity has to be earned by the student and "results from a commitment by the individual to seek understanding and purpose in any activity," meaning that it cannot be just attributed or thrust on a particular activity (2009, p. 400). In other words, the student has to put in the effort to participate in the process of authenticating any activity, and authenticity is the product of this process. That is to say, authenticity is not achieved by just "engaging with a context or materials that someone else had judged to be 'authentic'" (p. 400). Implications of this idea are found in van Lier (1996):

> One cannot say that any particular teaching method is more likely to promote authenticity than any other, regardless of whether or not it promotes the use of "genuine" materials. Rather, the people in the setting, each and every one individually for himself or herself, as well as in negotiation with one another, authenticate the settings and the actions in it. (p. 128)

Therefore, merely introducing APL into a science classroom does not translate the students' experience into "an authentic experience of science" (Osborne, 2009, p. 400); what the students seek is a learning experience that is authentic. What is an authentic learning experience? It is one in which the students' knowledge is increased; their level of comprehension is heightened; and they are rewarded with "a feeling of success, revelation, and meaning" (p. 400). Osborne admitted that much of school science does not feel authentic to students. However, students may or may not share the view (Osborne, 2009, p. 401) of Yarden et al., (2009) that APL materials, which are similar to texts used by scientists, "will be perceived as authentic..."

The Response to Osborne. Norris et al. (2009a) responded to the three issues raised by Osborne (2009) with respect to the assumptions in Norris et al.'s work. The paper (2009a) agreed that it is very important for citizens to be adept at reading media reports, but noted that media reports do not adequately depict the epistemology of science. Consequently, reading media reports cannot replace learning to read APL, but the two scientific genres can be considered to be complementary, each having its own lessons to offer the novice reader. The problem is to find the time in the curriculum to accommodate both genres so that there is no trade-off between "knowledge of the epistemology of science and the ability to read media reports," because both are important (Norris et al., 2009a, p. 406).

The Norris et al. (2009a) study found that the costs to gain entry into the subject matter discussed in various APL examples they cited were not too high. The authors remarked that on account of the fact that science is cumulative, if all the knowledge in the field has to be learned before engaging with the APL, then the "entry cost" would be too high, even for aspiring scientists. From their experience using APL in high schools and reading about its usage in various disciplines of science, Norris et al. felt confident that students can not only read but understand "Science on the cutting edge" using APL (2009, p. 406).

A second assumption in Norris et al. (2009a) is that reading is perceived as "an act of inquiry" (p. 406). Although Osborne conceded that "… reading … is definitely a process of inquiring into meaning," he finds it difficult to reconcile how reading itself "can be seen *as inquiry*" (2009a, p. 407). There are a number of arguments to tease out here. The work of Phillips and Norris (2009) viewed reading as "a process of inquiring into meaning" and the authors came to the conclusion that "reading *is* inquiry" (Norris et al., 2009a, p. 407). By this they did not mean that reading constitutes the whole of scientific inquiry, rather, that it is one important and often overlooked part of scientific inquiry. In an earlier paper (Norris & Phillips, 2008), these authors stated that reading is "as much a part of scientific inquiry as are observation, measurement, and calculation" (p. 407). In stating that "inquiry is a process requiring a multiplicity of actions," of which reading is but one, Norris et al. (2009a, p. 407) agreed with Osborne (2009). However, Norris and Phillips insisted that "reading by itself is inquiry" as clearly argued in their earlier paper (2008, p. 407).

The third assumption made by Norris et al. in their work (2009a) is that the usage of APL can be justified by its authenticity. The authors agreed with the viewpoint in Osborne (2009) that being authentic to science in some specific, defined way can never be used as a justification for adopting APL as an educational practice. Norris et al. added that there are so many differences between the context of research science and the context of school science that any claims of authenticity must be assessed cautiously. They pointed out that authenticity does not guarantee that the introduction of APL into the classroom will be effective. They also cautioned that "learning by inquiry is not a built-in characteristic of APL" (Norris et al., 2009a, p. 407). It is only when suitable classroom interactions with the APL text are augmented with teachers' pedagogic content knowledge (PCK), that the ground is fertile enough for learning by inquiry to sprout and emerge (Norris et al., 2009a). APL does have the potential to achieve what

Osborne described as the “primary goal of science education,” which is to furnish students with ways in which scientific theories about the world around them are justified (Osborne, 2009, p. 400). Since all else that has been tried so far to attain this goal has failed, APL deserves to be given a fair chance (Norris et al., 2009a). My study attempts to provide APL this chance, and in order to make this effort fair, it is important to consider the factors which influence how first-year undergraduate students learn.

Factors Influencing How First-Year Undergraduate Students Learn

Those who have had experience in dealing with first-year undergraduate students know that these students have a variety of adjustments to make in their first year. They are faced with a number of challenges: they have to develop academically and intellectually; they have to maintain personal relationships; they are required to develop an identity; it is essential that they maintain personal health; and it is imperative that they develop an integrated philosophy of life (Carnegie Mellon University, 1996). These challenges can be quite daunting especially because most of these students are attempting to be successful devoid of the support structures they were used to in high school. Many of the first-year students, therefore, are doing things which in hindsight may seem minor, but appear to be quite intimidating when they are doing them for the first time.

Some students handle this transition with ease, but many others face problems while adjusting, especially during the first few weeks of the first semester. The first-year students who face problems may be realizing that their old strategies do not work, and new support systems are not forthcoming. Some with problems may be reluctant to ask for help and, further, may not know where to get it. These challenges are particularly unnerving when students are placed in large lecture courses where they feel anonymous and insignificant in a learning environment they are not used to. These common challenges are further exacerbated when the students are international students, whose first language is not English.

Many faculty members and teaching assistants teaching first-year undergraduates report that students in introductory science courses have little prior knowledge of the field or topics covered. To make matters worse, what these freshmen do know "is poorly organized, incomplete or simply inaccurate" (Carnegie Mellon University, 1996, p. 5). Several of these first-year students report that they succeeded in high school without exerting themselves very much and are very surprised that they have to work so rigorously in postsecondary institutions.

In addition, the demographics of the students who go to college have changed in a number of significant ways within the last two decades. Yet the predominant "structure" of the first college year has been preserved similar to what was designed "for a population of white, middle or upper class males," who constituted the majority of college going students two decades ago. Consequently, among today's diverse freshmen population, there is "a serious lack of institutional fit, not of their making" (Barefoot, 2000, p. 13). In many colleges the first-year classrooms are the institutions' "cash cow" (2000, p. 17). Learning goals in these classrooms are compromised for cost-effectiveness. For example, teaching assistants, or other low-cost faculty, teach as many students as possible in large classes, where the only viable mode of instruction is lecturing. Lectures tend to be boring and typically student attention increases from the start up to ten minutes into the lecture and declines after that (Goss & Bernstein, 2005). I think that the lecture should be complemented with more active learning strategies like employing visual aids such as demonstrations, and the asking of meaningful questions to propel student interest forwards as the lecture progresses.

There had been a lot of debate, over the last two decades that first-year instructors do not cater towards their students' attention span. Since a lot of college courses are lecture-based, there is usually a teacher talking to a large number of students. This permits students to be inattentive and to be on their phones during the span of the class. Although it is generally believed that first-year undergraduate students have an attention span of 10 to 15 minutes in class, a 2007 literature review by Wilson and Korn found that this is not always the case, and confirmed that there is no evidence to support this belief. When Wilson and Korn (2007) examined the references supporting this very precise attention span estimate, they found that these studies were based on inaccurate measures of student attention or engagement. For example, a study found that note-taking declines from the beginning to the end of the lecture. Researchers were interpreting this as the evidence of the student's attention span over the course of the class period. This is not the case: the note-taking could have declined for a host of other reasons (such as tired fingers) other than the attention span. In fact there is no conclusive evidence to show that first-year undergraduates have an attention span that generally lasts 10 to 15 minutes.

It is unfortunate however, that too many students begin a first degree and leave without completing it. For example, in the United States, only half the students (51%) who were enrolled at four institutions in 1995-96, completed the degree within six years at the various institutions where they were registered (Kuh, Cruce, Shoup, Kinzie, & Gonyea, 2008). Kuh et al. (2008) have carried out research showing how engaging students in educationally purposeful activities can promote desired outcomes such as the grades these freshmen obtained in the first year and on their retention in the second year of college at the same postsecondary institution (p. 555). This compensatory effect of student engagement suggests that postsecondary institutions should make further attempts to channel student energy into "educationally effective activities" (p. 555). This is especially crucial for those who begin postsecondary education with two or more “risk” factors. Some of these risk factors include being underprepared academically, or being the pioneers in their families to go to college, or coming from low income backgrounds” (p. 555).

In universities around the world, courses have been conceived just for first-year undergraduate students with the goal of directing “student energy” and effort “toward educationally effective activities” (Kuh et al., 2008, p. 555). One such course offered at the University of Alberta is Science 100, which is a multi-disciplinary course designed to give first-year undergraduate students a strong academic education in all aspects of science (University of Alberta, n.d., para 1). It is a one full year course that explores and teaches the concepts and foundations of seven scientific disciplines in a unified and integrated manner. This offers first-year undergraduates a unique opportunity to explore different sciences, see the connections between them, and apply that knowledge to solve complex problems (University of Alberta, n.d., para 2). For students who like science but are not sure which discipline to choose, this course provides the ideal chance to explore seven disciplines in this subject, and to appreciate, as never before, the fascinating connections between them.

Science 100 allows first-year undergraduates to be in small classes, giving them more opportunity to interact with teaching staff and fellow students. The philosophy of this course is that the last two decades have witnessed unprecedented discoveries across all disciplines of science. Many of these have been spurred by collaboration between fields of science which have traditionally remained distinct. In order to tackle the local and global issues facing humanity today, expertise from numerous different fields will be required. Science 100 celebrates this new way of thinking by applying an interdisciplinary approach to science education. This is done by incorporating three main principles: (1) an integrated science education; (2) discovery-based hands-on learning; and (3) collaboration and community engagement. The role of collaboration and community engagement is also emphasized by a most recent model of scientific reasoning which treats science as a social practice (Osborne, 2013).

Science as a Social Practice

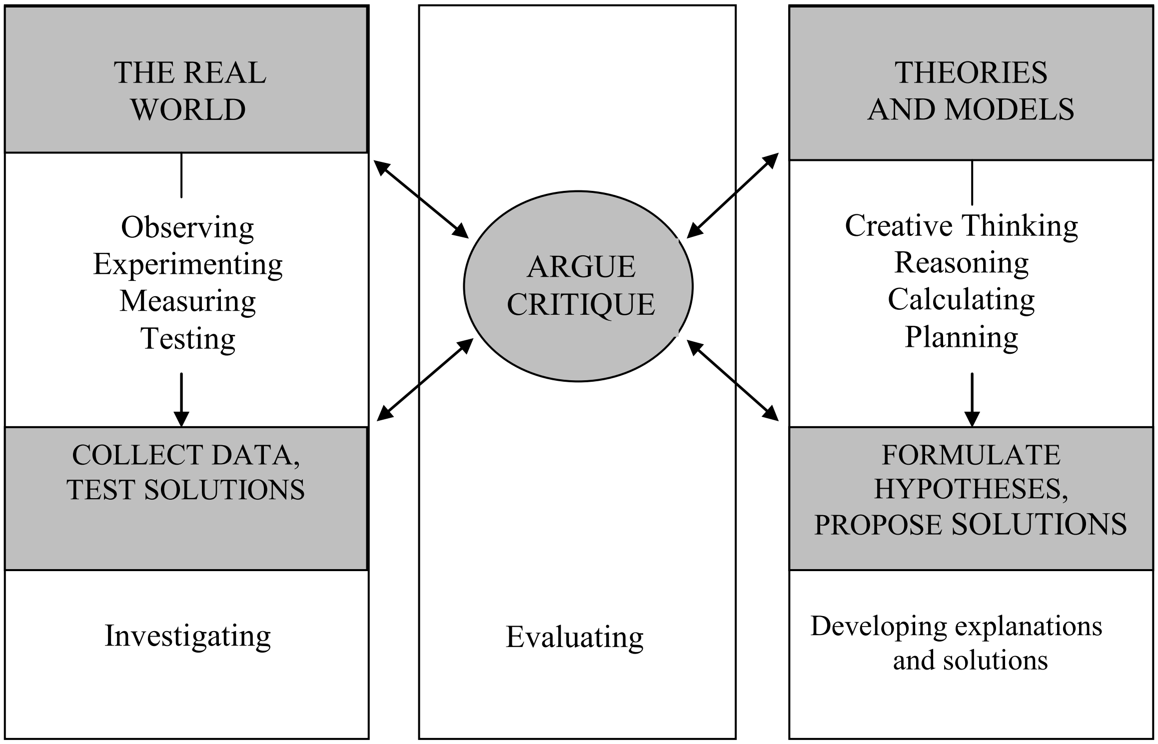
Science is something which is done. It involves some kind of activity on the part of the scientist. As a result of this activity, knowledge is created. The role of the teacher of science is to tell a story about how that knowledge was created. In other words, how does one address the question of "*how we know what we know*" (Osborne, 2011, p. 93) in science? One of the major stumbling-blocks used to explain this epistemology of science has been the myth of the "scientific method." Why then is this idea of the "scientific method" so flawed? I would say this concept has two major shortcomings. First it seems to suggest that there is one universal method shared equally by the enormously diverse disciplines of science ranging from, for example, the theoretical cosmologist working in the office to the ecologist working in the field (Osborne, 2011). Second, it is unable to accommodate the great diversity of methods employed by what is known as science - with all its interdisciplinarity - in the 21st century. As described in the brief history of science that I have traced in this chapter, there have been several attempts to debunk this myth of the "scientific method." The most recent effort was made by the U.S. National Research Council in its report titled "A Framework for K-12 Science Education" (NRC, 2011), which puts the emphasis on scientific practices.

The notion of science as an "ensemble of specific practices" has emerged in the last 50 years (Osborne, 2011). According to Osborne (2011) "we can provide a better science education, and a more accurate picture of science, by presenting science as a set of distinct social practices that, to a greater or lesser degree, are shared by all scientists" (p. 94). Being focused on a set of distinct practices will enable the scientific community to be clearer about their educational objectives. The genesis of new scientific knowledge then, is not brought about by the mere interaction of scientists with the material world, but is a process influenced by the community within which scientists do their research work (Osborne, 2011). Moreover when practices such as argumentation are emphasized, students are compelled to reflect on their practice and their understanding of science. When social practices are put to the fore, students will have to make their thinking visible and thereby expose their misconceptions, which could be readily picked up and addressed by their instructors. The overall effect would be an improvement in the learning of science (Osborne, 2010).

Osborne’s model (2011, 2013), which captures the overall activity of science, shows scientists as engaging in the three overarching practices of: (1) investigating; (2) developing explanations and solutions; and (3) evaluating (See Figure 3). On the left of this figure, scientists are working in the investigational space, where they design experiments, as well as collect, analyse, and interpret data (2011). On the right, the scientists are engaged in developing hypotheses and constructing explanations, in other words theorising about the world. Such theorising is less common in the teaching of science at present (2011). In the middle of the model, scientists are engaged in argument and critique, while they evaluate the validity and reliability of their data; contrast what they have obtained as data with their theoretical predictions; and spot the flaws in their own thinking and in the ideas of others working in their field (2011). Such evaluating activities are even less common in the teaching of science today.

Figure 3

A Model of the Practices of Science



Adapted from Osborne (2011, 2013)

According to Osborne (2013) all scientific activity can be described as a set of eight practices: (1) asking questions and defining problems; (2) developing and using models; (3) planning and carrying out investigations; (4) analysing and interpreting data; (5) using mathematical tools; (6) constructing explanations; (7) engaging in argument from evidence; and (8) obtaining, evaluating and communicating information (p. 95). I think each of these practices would require some, or at least one, Concept of Evidence to fulfil its function accurately, effectively, and with inherent validity. Although the Gott and Mashiter model (1991) depicted substantive and procedural knowledge and how it was used in problem-solving, Osborne's model

(2013) represents what happens as people use that knowledge in the practices of science. It takes into consideration what happens as knowledge interacts with people individually but even more so what occurs when scientific knowledge is practised with others in society.

The expectation the world holds as the outcome for science education in the 21st century is an increasing focus on higher order thinking skills such as synthesis, analysis, and evaluation. School science, however, still has a prevalence of the lower level cognitive skills, recall in particular. Osborne (2013) maintained that the failure to transform science education to meet the emergent needs of the 21st century is mainly due to "the lack of a good model of scientific reasoning" (p. 265) and the absence of "a body of expertise about how to assess such higher order cognitive competencies" (p. 265). Osborne (2013) has contributed to filling this void for a model of scientific reasoning by proposing the "Model of the Practices of Science."

Summary

The different methods used to acquire knowledge in science resulted in the various trends and themes observed in the teaching and learning of this subject over the ages. The history of school science education has oscillated across national boundaries like a pendulum, swinging from an emphasis on substantive knowledge (e.g., in the 1900s) to a stress on procedural knowledge (e.g., in the 1950s and the 1960s). A model of science was developed by Gott et al. that differentiated knowledge into substantive (i.e., content) knowledge and procedural (i.e., the "how to do" of science) knowledge. These authors defined Concepts of Evidence as the building blocks of procedural knowledge. There is a dynamic interaction between substantive knowledge and procedural knowledge which has been illustrated by one detailed example from first-year undergraduate chemistry classes.

Considering concepts of knowledge as a framework for building procedural knowledge has serious implications for the pedagogy and the teaching of science. Students come to class with certain erroneous conceptions about the topic being taught. These erroneous conceptions have to be identified, addressed, and corrected to bring about conceptual change and deepen the learning of students.

By undertaking this study one hopes to find that Concepts of Evidence can not only be taught or learnt but that pedagogy itself can have a significant impact on the extent of the teaching and learning processes. One would expect the teaching intervention to improve the students' learning of the Concepts of Evidence, as well as be perceived as a very useful exercise by the majority of students. One would anticipate the teaching intervention to be effective in bringing about conceptual change in the learning of the substantive knowledge and the Concepts of Evidence included in the Adapted Primary literature (APL).

One cannot expect the students to be strongly opinionated about introducing Concepts of Evidence in undergraduate chemistry courses because they themselves do not know what it is like to go through the curriculum with a prior understanding of these concepts. However, one would anticipate all instructors to have an opinion about the introduction of Concepts of Evidence in their teaching of chemistry at the first-year undergraduate level, and if they do not then they could be requested to shed light on their reservations.

Scientists communicate the results of their research through journal articles referred to as primary literature. From reading the literature, one would anticipate that most students would initially struggle through the APL, but those who persist would find it rewarding, because of the scaffolding provided. However, one would expect various sections of the APL to evoke diverse challenges for individual students. Based on constructivist principles I would expect the students to find the teaching intervention very educative and effective. Many students might find the post-test easier, and may perform better on it, based on how much they had learnt from the teaching intervention; and how effective the teaching intervention had been not only to inform them but to correct their misconceptions. One would expect the students to clarify some of their misconceptions through the interviews as well. The existence of a single scientific method which can be applied to all scientific activities is a myth. Osborne insisted that the failure of science education to meet the needs of the 21st century is the absence of a good model of scientific reasoning and he suggested the “Model of the Practices of Science” to fill this void.

CHAPTER THREE

RESEARCH DESIGN: METHODS AND METHODOLOGY

Prelude: The Growing

The rice plants growing in the paddy fields are swayed by the gentle breeze. These plants are provided with plenty of water and sunlight. The seedlings grow into green, verdant, one- to one- and-a-half meter high plants. The lush green of the paddy fields is one of the most striking and memorable features of Kerala’s landscape. However, if the rich, rice grains of the paddy fields are to be revealed these rice plants have to be further examined. So, too, if the results of this research project are to be revealed, the research design has to be further scrutinized.

Methodology

Individual and Socio-political Construction of Knowledge.

von Glasersfeld (1991), who had very great influence in the contemporary international science and math education communities of his time, contended that students construct their own individual understanding. Teachers cannot assume that all students have the same understanding or that their understanding is similar to the teacher's. The constructivist elements in John Dewey's (1969) epistemology (Third axis in Figure 4) have social and pedagogic ramifications. According to Dewey, as cited in Phillips (1995), the knower is an "actor," not a "spectator" (p. 6). This is true of both the evidence survey and the interviews that were conducted in this study. Participants in the evidence survey and in the interviews could not just be spectators, but had to be actors because they had to think about the questions and divulge their own understandings of the Concepts of Evidence and the substantive knowledge in the APL.

According to the tenets of constructivism, when humans act as creators of knowledge there are two options possible, as shown in Figure 4: humans can be responsible for the socio-political construction of knowledge, or for the individual creation of knowledge. In my research project both these options can be identified. When individual students have to understand what Concepts of Evidence are, and how they function as the building blocks of procedural knowledge, they have to grapple with these ideas in their minds and construct schemata for themselves. They can then fall back on these schemata when they meet new Concepts of Evidence or are faced with the old Concepts of Evidence in novel contexts. This is referred to as the "Individual Creation of Knowledge," and is the second option under "Humans the Creators" of knowledge (Phillips, 1995).

The compendium of 21 Concepts of Evidence, on the other hand, is knowledge that was socio-politically constructed by the research of Dr. Richard Gott and his colleagues at Durham University in the United Kingdom. This research, extending over two decades, has involved Concepts of Evidence which have been tested and implemented in schools in the UK. Some of the work done by Dr. Gott and his colleagues has been incorporated into the science curriculum in the K-12 arena in the UK. This involves not only Dr. Gott and his research group, but science curriculum developers, schools, teachers, and students, making it therefore, a socio-political construction of knowledge.

The curriculum is a designed experience intended to promote and direct learning. When we reflect upon fundamental issues of the curriculum such as, "What should be studied?" "Why?" "By whom?" and "In what ways?" there is an inter-play among the major components of education: subject matter; teaching; learning; the larger social, political and economic contexts; and the immediate instructional environment (Gorlewski, Gorlewski & Ramming, 2012). The science curriculum is socially constructed because people who have learnt both science and its pedagogy, and those who have an aptitude for this subject, construct the curriculum. This curriculum is then taught in the classroom which is a social setting. As such, curriculum is a social construction, produced as a result of a set of decisions. Because it is socially constructed, the curriculum is subjected to all the pressures and complications of the diverse communities that comprise schools and other institutions of higher education.

The curriculum has also got a political dimension. Research in science education has shown that science learning is related to the social, economic and cultural status of students. In other words, students who face inequalities in their daily life, such as social exclusion or poverty have fewer opportunities to learn science (Aikenhead, 2006; Brickhouse & Kittleson, 2006). In addition, until Indigenous knowledge becomes part of the science curriculum in institutions of higher education, we will not be able to remediate the learning difficulties of non-westerners in learning science (Hewson, 2015). According to some science educators (Hodson, 2003), “it is time for a science curriculum oriented towards socio-political action” (p. 645). Such a science curriculum would enable young citizens to examine critically the values held by the society of today, and ask “what can and should be changed in order to achieve a more socially just democracy and to ensure more environmentally sustainable lifestyles?” (p. 654). Such a view of science education would be “overtly and unashamedly political” (p. 654).

Constructivism as a Theoretical Framework

Before I delve into what the literature has to say about constructivism, I would like to describe upfront what I think of this learning theory and how it is reflected in my study. Constructivism, I think is based on observation, and is a scientific study about how people learn. I believe that most people construct knowledge through their own experiences and through reflecting on those experiences. We are, therefore, active creators of our own knowledge. Based on how I think of constructivism, I have initially provided students in my study with a chance to interact with new knowledge through reading the Adapted Primary Literature (APL). Their knowledge was further developed through the pre-test, teaching intervention, post-test and culminated in the interview. As I progressed through the study, however, I was propelled by optimism, tinged with a degree of what might appear to other readers as positivism and realism – a compelling desire to represent the experience of how knowledge was acquired by these students and their instructors accurately, in a manner that was authentic and true to life. Let us now examine what the literature has to say about constructivism.

Constructivism has been a very influential paradigm in the realm of science and mathematics education. It was in the 1980s that constructivism was initially developed as a theory of learning. According to constructivism, different individuals have their own methods of constructing knowledge, and learning occurs through the active involvement of the learner in this process of construction (Brill et al., 2004). The constructivist approach emphasises the fact that if learning is to occur, one must reconstruct knowledge, and appropriate it as one's own during the learning process. It therefore implies that it is not only impossible for a teacher to "teach a body of knowledge by direct transmission," but it is imperative that the learner be involved simultaneously in "reconstructing the meaning personally" (Millar, 1989, p. 592). Schwandt (1998) defined constructivism as follows:

> Constructivism means that human beings do not find or discover knowledge so much as construct or make it. We invent concepts, models, and schemes to make sense of experience and, further, we continually test and modify these constructions in the light of new experience. (p. 237)

As stated above, constructivism implies that scientific knowledge is comprised of an understanding of "consensually agreed models, theories and knowledge structures" (Çalik, Kolomuç & Karagölge, 2010, p. 422; Gilbert & Boulter, 1991; Wertsch, 1991).

A second, key tenet of constructivism is that learning is brought about by the interaction that occurs between pre-existing knowledge and new knowledge (Driver, 1981; Driver & Easley, 1978; Taylor & Coll, 1997). When this interaction occurs, students may generate concepts that are different from the consensually accepted views held by the scientific community. Such concepts are called by terms such as misconceptions and alternate conceptions and are discussed in Chapter Five under the heading "Conceptual Change."

A third primary tenet of constructivism is that prior knowledge is important, and it plays a significant role in further learning (Çostu & Ayas, 2005; Driver & Easley, 1978). Numerous research studies in science education demonstrate, however, that these students' misconceptions are "robust and highly resistant to change" (Çalik et al., 2010, p. 422) when traditional teaching methods are used (Banerjee, 1995; Çalik & Ayas, 2005; Coll & Treagust, 2001). Therefore, there have been calls to use pedagogies which are constructivist-based to overcome these students' misconceptions (Coll & Treagust, 2001).

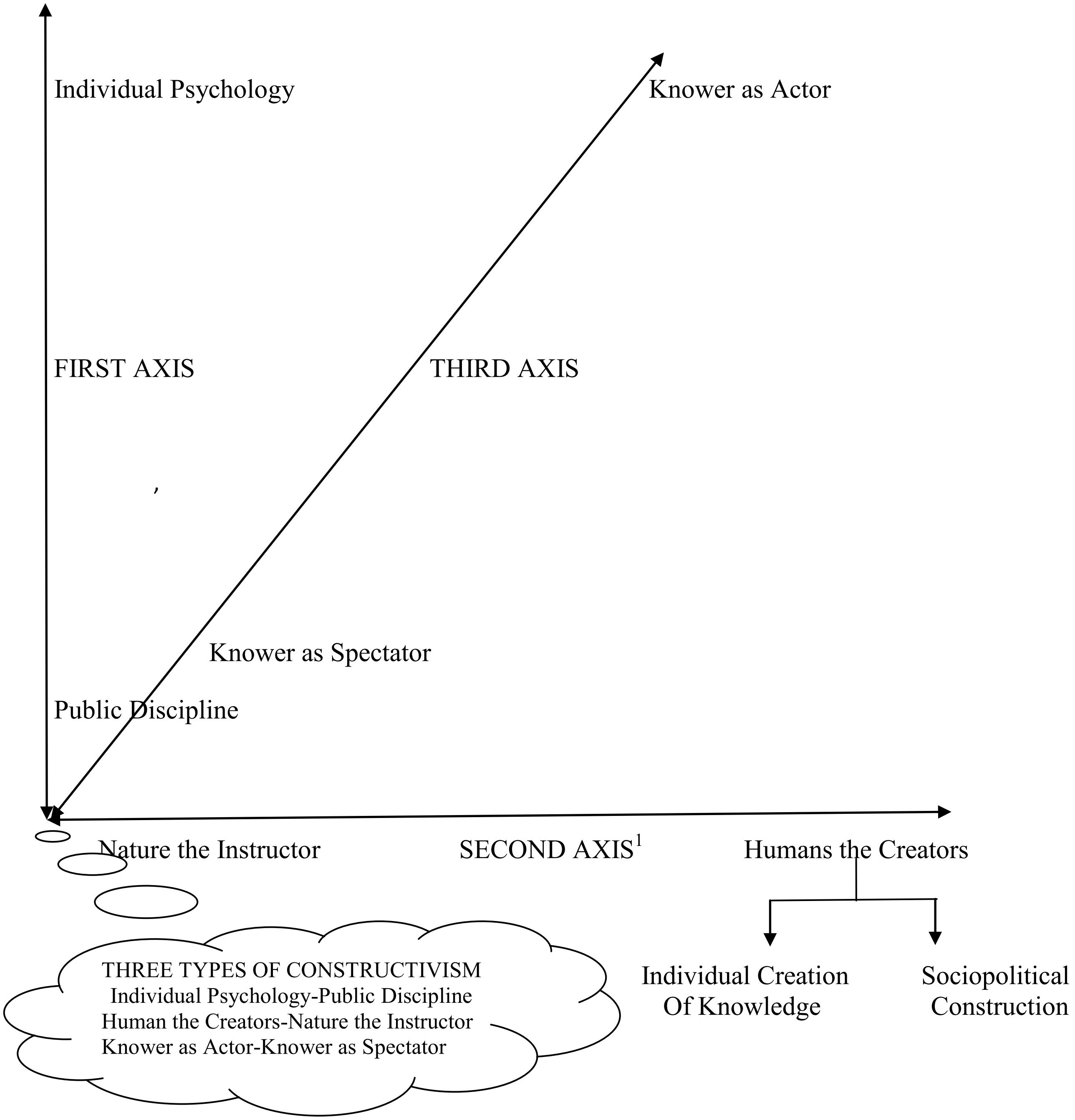
There are various forms/sects of constructivism and there is a plethora of science education literature on each of them. Each form is complex as it represents not a "single issue" position but addresses a number of deep problems both explicitly and implicitly. Because of their complexity, Phillips (1995) introduced a novel way of viewing the various forms of constructivism. According to Phillips, the main constructivist views can be placed along three different dimensions or axes, which I have drawn up and constructed as Figure 4. Each axis represents a spectrum of views. For example, the first axis concerned with how learners go about constructing knowledge, ranges from Individual Psychology where learners use their own cognitive apparatuses, to the other end of the spectrum, where the construction of human knowledge in general leads to the formation of Public Disciplines, such as the feminist epistemologies. The three axes proposed by Phillips (1995) and the spectrum of constructivist views they represent are outlined in Figure 4.

Theoretical frameworks are associated with certain basic assumptions, ways of thinking and methodology, which are shared by members of a particular discipline or group. Such a cognitive framework is referred to as a research paradigm. Inquiry paradigms such as constructivism clearly stipulate for researchers, what it is that they are about and what exactly falls within and what lies outside the limits of legitimate inquiry. The fundamental beliefs that define constructivism can be summarized by the responses given to three fundamental questions – the ontological question, the epistemological question and the methodological question.

The ontological question in constructivism deals with the "nature of reality" (Denzin &

Figure 4

THE THREE OF DIMENSIONS/AXES OF CONSTRUCTIVIST IDEAS



[1] Adapted from Phillips (1995)

Lincoln, 1998) and what can be known about it. If the world is assumed to be "real," then what one can enquire about this "real" world is "how things really are" and "how things really work" (p. 201). This means that only questions that deal with "real" existence and "real" action can be entertained. Consequently, questions related to other matters such as aesthetics or moral significance are inadmissible and precluded from "the realm of legitimate scientific inquiry" (p. 201). Since the theoretical framework that I have employed in my study is constructivism, the ontological question in this paradigm centers on relativism which promotes the construction of specific, local realities (Denzin & Lincoln, 1998). For relativists, reality is perceived as intangible mental constructions made from social experiences. The form and content of these constructions are determined by the individuals or groups who created them. Therefore, these mental constructions and the realities associated with them, can be altered and are subject to change. Consequently, these mental constructions cannot be classified as more or less "true" in any absolute sense of the word, but can only be envisaged as more or less "informed" or more or less "sophisticated" (p. 209).

The epistemological question, on the other hand, explores the nature of the relationship between the knower or prospective knower and what can be known. One cannot postulate just any relationship, but a "real" reality is assumed. In this case the knower must be objectively detached and must value freedom greatly to be able to discover/unpack "how things really are," and "how things really work." The investigator and the object which is investigated are linked interactively so that it is only as the investigation proceeds that the "findings" are *literally created.* The distinction that conventionally exists between ontology and epistemology gradually disappears.

The methodological question examines how the inquirer proceeds to find what he or she believes can be known. It is not appropriate to use just *any* methodology. For example, if a "real" reality is being studied by an "objective" inquirer, it is important to have mandates in place to control all possible confounding factors, irrespective of whether the methods used are qualitative (e.g. observational) or quantitative (e.g. analysis of covariance). Since these social constructions are personal and variable, they can only be elicited and refined through prolonged interactions between investigator and respondents. The varying constructions are interpreted using traditional hermeneutical techniques, and are "compared and contrasted through dialectical interchanges" (p. 207). The final objective is to arrive at a "consensus construction" (p. 207) which is more "informed" and more "sophisticated" (p. 207) than any of the preceding constructions.

Constructivist Pedagogy

Constructivists focussed on learners, reiterating that they had the onus for their own learning (Novak & Gowin, 1984; Osborne & Wittrock, 1985; Pope, 1985; White, 1988). Constructivists were of the opinion that it was in the mind of learners that "new meanings had to be formulated and understood," (Osborne, 1996a, p. 63) and this could be realized only when they participated actively in the learning process. Consequently, the constructivist pedagogy got involved in creating "a program of activities" (p. 63) from which knowledge could be acquired or generated. These activities required individuals to employ their reasoning and articulate their thoughts while solving a set of structured exercises, which often included working in groups. One of the most renowned "structured strategies" (p. 64) involved using the "predict-observe-explain" (p. 64) sequences.

There has been a call, in the last two decades, to change the way we teach science from the more teacher-centered delivery to a more student-engaging method that fundamentally recognizes the way people learn (National Research Council, 2000). The human mind can only assimilate at a certain controlled rate and comprehend only a limited amount of new information (Eberlein, Kampmeier, Minderhout, Moog, Platt, Varma-Nelson and White, 2008). Transferring knowledge more or less directly from teacher to student by using the common "teaching by telling" method has been found to be ineffective for many students (Eberlein, Kampmeier, Minderhout, Moog, Platt, Varma-Nelson and White, 2008, p. 263). In science courses, as the cognitive load increases, it becomes increasingly important that students are involved actively in their own learning. Students must construct for themselves a "workable understanding of sophisticated concepts" (p. 263), and they must be encouraged to develop their own higher-order thinking skills. The main active learning pedagogies employed in science Education are: team problem-based learning; process-oriented guided inquiry learning; and peer-led team learning.

Each of the three active learning pedagogies mentioned above is based on the tenets of *social constructivism*. In science education, constructivism maintains that "knowledge is constructed in the mind of the learner by the learner" (Eberlein, Kampmeier, Minderhout, Moog, Platt, Varma-Nelson and White, 2008, p. 262; Bodner, 1986). According to social constructivism, this building of knowledge is brought about by cooperative social interactions. Social constructivism also implies that teaching by only delivering lectures, fails to exploit fully the unique characteristic of the educational milieu today, in which students are prepared to make a mutual effort to master both course content and the pedagogy required to accomplish this most effectively.

Since teachers in higher education, have not experienced any of these active learning pedagogies, there is a significant barrier to accepting any of these unfamiliar methodologies. Consequently, there is a gap between what is now known about the process of learning through science education research, and the way that science is often taught at present (Anderson, 2007).

Although some faculty may recognize the merit of active-learning methods, there is an activation energy required to try something new, and a fear of the consequences of failure that must be overcome to plunge into the unknown. This is a source of major concern for educators interested in reforming science education by adopting appropriate active learning strategies. Yet in science education, it has been shown that "faculty who embrace the need for instructional change and conquer the initial steep learning curve, find the experience transformative" (Eberlein, Kampmeier, Minderhout, Moog, Platt, Varma-Nelson and White, 2008, p. 270).

The success of constructivism has been that it has produced a large and significant amount of empirical data which has played an important role in helping us understand the difficulties that students experience in the learning of science. This knowledge has enabled teachers to develop some innovative methods for teaching science, and has prompted a greater awareness of the scientific thinking of students, especially children. However, constructivist theory has its weaknesses as well, which will be discussed below.

Shortcomings/Failings of Constructivism

Constructivism is an umbrella term covering a spectrum of views, as depicted earlier by the three axes of constructivist ideas. The major shortcoming of constructivism is that it has failed to distinguish the manner in which new knowledge is created from the way old knowledge is learned, based on the erroneous assumption that the two processes are one and the same. In my study, the Concepts of Evidence can have two major functions: (1) they can supply the procedural knowledge required, in the different stages of setting up and carrying out a laboratory experiment. In this instance, the students are reproducing established scientific knowledge; or (2) Concepts of Evidence can also be employed to provide the procedural knowledge required in the different stages of setting up a research project. Here teachers or students are pushing scientific knowledge to its limits and probing into the realm of the unknown. A constructivist methodology does not distinguish between the two major functions, (1) and (2), described above because it makes a false connection between the "manner in which new scientific knowledge is created and the manner in which existing scientific knowledge is learned" (Osborne, 1996a, p. 54).

The consequence of this failure is that constructivism offers no guidelines as to how to adjudicate theories, and provides no mechanism to establish the superiority or inferiority of one theory over another (Osborne, 1996a). Consequently, it is not able to offer any guidance for the organization and sequencing of content in the science curriculum especially when this involves adjudicating between theories. "Constructivism singularly fails to elaborate any mechanism by which one theory may be considered more 'viable' than another" (p. 58). The fundamental epistemic problem that constructivism faces, therefore, is that it does not "articulate a mechanism to establish how one theory may be considered to be "better" than another" (p. 59). As a result constructivism has failed to be of value in resolving issues of pedagogy in science education, especially those linked to the curriculum (Osborne, 1996a).

For example, in my study, as students grew in their understanding of the content of the APL, and the Concepts of Evidence nested therein, some of their thinking underwent certain "conceptual changes." Constructivism sheds no light on what constitutes such conceptual change, and fails to specify any mechanism for it. There is no role, for example, for analogy and metaphor, which often are not only utilized for expanding our thinking and ideas, but for "reorganizing our internal symbolic representation" (Osborne, 1996a, p. 65). Despite the fact that these operational processes which play a role in the construction of knowledge are "complex, poorly understood and nontrivial" (p. 66), constructivism fails to clarify them any further.

Research Design

In my study, I explored the understanding of Concepts of Evidence held by various groups of students and teachers. The "understanding" of Concepts of Evidence is defined as having a mental grasp of the ideas about evidence conveyed by these concepts. In order to examine in depth the understanding of Concepts of Evidence held by first-year undergraduate students and their instructors, my study needed a research design that could trace the dynamic interrelationship between the Adapted Primary Literature (APL) and the teaching intervention. To this end, I used the theoretical frameworks of constructivism, which was discussed in more detail earlier in this chapter. I adopted a qualitative approach in this research project to address my research questions. However, I collected data from interviews of students and their instructors; and an evidence survey administered only to my student participants.

Coming to the Study

As mentioned earlier, my parents were the reason that I took up teaching as a career. However, what I have not mentioned is that my parents saw their teaching careers in secondary schools as the will of God for their lives. They believed they had a calling to teach, educate, support, and guide children in Africa, paying special attention to those who were either under-privileged or from dysfunctional homes. Following in my parents footsteps, I started my teaching career at Njase Girls' Secondary School in Choma, a small town in Zambia. Though I enjoyed the acceptance, warmth and the Christian foundation of this school, I felt I needed to further my education. I got married, moved to London, England, with my husband, and pursued a Master of Science in Biochemistry.

After completing my Master of Science degree, we moved to the mountain Kingdom of

Lesotho in southern Africa, where I had some of my most memorable teaching experiences. I first taught biology, chemistry and physical science at a girls' high school run by the Anglican Church in the capital city, Maseru. When I first started teaching at this school, the laboratory was always locked and never used. I got the key and opened up the lab to find it reasonably furnished but in a dirty mess. The chemical store leading out of the lab had a broken window and an eerie look. The chemicals in bottles and tins were covered with dust and cobwebs. I got the glass window pane replaced, and the lab dusted and swept. I was given a lab technician and between the two of us, we got the lab cleaned, renovated, and functional. I also started to teach the Grade 12 (i.e., form five according to the GCSE system adopted in Lesotho) students on Saturday mornings from 9:00 am to noon, clearing any doubts they had or going over material that they had not followed in my classes during that week. We would then break for lunch, but the girls would eat quickly and return to the lab to work on their own individual projects for the national science fair. In the two years that I was there, our school entries always won the top prizes at the annual, national science fairs organized by the British Council in Maseru, the capital of Lesotho. I realized how important procedural knowledge was in planning and implementing these science fair projects, and was made crucially aware of how ignorant high school students were of this form of knowledge. The science results of the General Certificate of Secondary Education exam (i.e., the Grade 12 equivalent) improved remarkably during the two years (especially in the second year) that I worked at this school.

I was then offered a position as lecturer in the chemistry department at the University of Lesotho, one of the oldest universities in southern Africa. I taught general chemistry to the first-year undergraduates, biochemistry to fourth year undergraduates, and analytical chemistry at the third year level. I had to both set up and run the analytical chemistry labs, which alerted me of the importance of procedural knowledge in working accurately and precisely in chemistry. As a family, we then moved to South Africa, where I worked at the University of the North West for six years. I taught inorganic and organic chemistry at various levels ranging from undergraduates to the BSc Honours level.

From South Africa, my immediate family and I emigrated to Edmonton in Alberta, Canada. I took up a position as a teaching assistant to teach the labs for Chemistry 101 (Introductory University Chemistry I), in the Chemistry Department at the University of Alberta. I observed that first-year undergraduate students often come to the laboratory with only a vague understanding of the experiments they are assigned to carry out. They carefully complied with the instructions, without thinking about what they were doing and why they were doing it. The main reason for this mechanical manner of performing experimental tasks, I think, is that students have not been introduced sufficiently to the theory that is the basis of procedural knowledge. Although procedural knowledge deals with the "how-to-do" of science, it has a knowledge base akin to substantive knowledge. Students must be taught this theory, and it is my belief that the best way to do this is to introduce them to the 21 categories of Concepts of Evidence which are the building-blocks of procedural knowledge.

My research thesis at the Masters' level enabled me to examine how Concepts of Evidence were represented in chemistry textbooks used across Canada, at the high school and first-year university undergraduate level. Though they were implied in the textbooks, these Concepts of Evidence were never addressed specifically or taught explicitly in the texts. I was convinced then that these Concepts of Evidence should be explicitly addressed if students are to appreciate their relevance in the "doing" of science. I also felt that the policy governing curriculum should be changed to include the explicit teaching of Concepts of Evidence at the high school level. It is with these underpinning beliefs that I embarked on doing research at the PhD level on the understanding of Concepts of Evidence held by first-year undergraduate students taking Chemistry 101 or its equivalent in various higher education institutions spread across the province of Alberta.

Methods

Initial Preparation

After searching through the literature intensely, I first identified a suitable piece of primary literature by selecting the paper, "Molecular Beacon Probes of Photodamage in Thymine and Uracil Oligonucleotides" by Yarasi, McConachie, and Loppnow (2005), from the *Journal of Photochemistry and Photobiology*. I was encouraged in my choice by a professor in the chemistry department at the University of Alberta, who is also the lead investigator in this work funded by the National Science and Engineering Research Council (NSERC). I carefully thought out and answered the five questions posed by Muench (2000), described in Chapter Two, before implementing the Adapted Primary Literature exercise in my pilot study. To convert this primary literature into Adapted Primary Literature (APL), I had to familiarize myself with DNA photochemistry. My major source of reference in this endeavour was *The CRC Handbook of Organic Photochemistry and Photobiology (1995).*I made a great effort to identify and resolve the practical and pedagogical issues involved to make this experience valuable for the students and instructors who participated.

Choice of Primary Literature.

The first primary literature journal article that I chose, therefore, was on the usage of molecular beacons to detect photodamage in thymine and uracil oligonucleotides. There were several reasons for this selection. The primary reason was that molecular beacons are a versatile tool used for various biotechnological applications. Their use as sequence-selective detectors of nucleic acids was a relatively novel application. Although molecular beacons can easily detect single-base mismatches, they have never been shown to be able to directly detect DNA or RNA damage, as demonstrated by Yarasi et al. (2005). This application of molecular beacons, therefore, seemed very worthwhile because they were being used to detect the early stages of damage to DNA or RNA in skin cells exposed to ultra-violet light. The damaged DNA could result in a mutation which could then become cancerous. Therefore, detecting such specific, minor changes early could be a very innovative, life-saving application. Secondly, the primary literature contained a lot of substantive knowledge which was clearly explained in an explicit fashion, and which could be tested. It also included a good repertoire of procedural knowledge because many Concepts of Evidence were included implicitly. Thirdly, the two kinds of knowledge – substantive and procedural – were placed in a dynamic relationship with each other, so that the understanding of one promoted the learning of the other and vice versa. Fourthly, it outlined each stage in the research activity with figures which showed the initial results obtained. These diagrams helped to understand clearly the sequence of the experiments and could be used as stand-alone units for testing procedural knowledge which was only implied. Fifthly, most experiments were performed in sets, using both types of oligonucleotides (i.e. dT_{17} made from 17 thymine nucleotides; and rU_{17} containing 17 uracil nucleotides) employed in this research activity. These experiments made measurements using two major experimental techniques: absorbance spectroscopy and fluorescence spectroscopy. Hence there were two sets of experiments for each stage of the analysis. This provided two sets of very comparable, but not identical data, differing only in the fact that one used thymine found in DNA while the other employed uracil which occurs in RNA. It was very valid, therefore, to use one set of results for the pre-test and the other set of comparable data for the post-test in my study.

In addition, I hasten to add that my own personal bias for biotechnological tools linked to an early detection of diseases may also have played a role in the selection of the primary literature. This is problematic in that some first-year students taking science courses may be interested in hard-core science disciplines such as physics, computer science and chemistry. These students may not be inclined to appreciate biotechnology techniques which require a well grounded understanding of genetics and the structure of DNA and RNA. Since all my participants are taking Chemistry 101, the hard-core science students in my study may be expecting to read an APL that is linked to some of the material they have covered in first-year chemistry. They may start this research activity with great interest but when they see that the APL requires a good understanding of genetics and biology, they may be disenchanted and lose their enthusiasm. One also has to choose a primary literature in which the Concepts of Evidence are either dealt with explicitly or implied. When choosing primary literature, therefore, one has to find a fine balance between appealing to student interest and selecting a paper that can promote the learning of Concepts of Evidence: this is a problematic but challenging endeavor.

As I delved further into the primary literature chosen, I noticed some technical short-comings in this journal article such as the number of trials done on each experiment being unspecified. I was told that this was because this was an initial paper by these authors on this topic. When I was asked for a later paper which continued this initial work, I was told the paper was still “in press,” and therefore, I was not able to use it. Since then, more papers have been published by these authors as a continuation of the primary literature that I chose (Yarasi et al., 2007; Wenger & Loppnow, 2004).

In addition, the units for the fluorescence intensity were measured as “arbitrary units” throughout this primary literature. I understood this to be a common practice among photochemists and photobiologists working with fluorescent probes. Apparently the fluorescence emitted by fluorescent probes varies so much from day to day that researchers working in these fields measure changes in fluorescence (ΔF) rather than actual fluorescence. ΔF is a relative quantity and therefore, can reasonably be assigned "arbitrary units." I converted this primary literature by Yarasi et al. (2005) into Adapted Primary Literature (APL) catering to first-year undergraduate students. This APL is included in Appendix A.

Construction of the APL from Primary Literature

I began writing the APL by dividing the page into two columns: the left column was the adapted form of the primary literature maintaining its canonical structure with difficult words underlined, while the right column housed the meanings of the underlined. I first converted the ``introduction`` into APL and immediately sent this to the chemistry professor (the chief author of the paper and an expert in the field of photochemistry and photobiology), and to a second professor whose expertise was in science education. When these two professors had read and evaluated the introduction of my APL, they sent it back to me to implement the corrections. This was so rigorous a process, done repeatedly and thoroughly, section by section, until the whole article was converted to APL, that I am quite certain that my understanding and interpretation of the primary literature were checked continuously and monitored constantly by two experts - one in chemistry and the other in science education.

Consequently, I was confident that my APL was a simplified but accurate representation of the primary literature. The role of my personal bias in writing the APL was kept in check by the expert knowledge held by the chemistry and science education professors. Taking this argument further I want to maintain that my APL does not have the same weaknesses as a newspaper article because it was checked and corrected by the author of the primary literature. Therefore, the biases in my APL are likely to be very similar to those in the primary literature: therein lies my confidence in the accuracy of my APL. My own personal bias may have had some effect on the writing of the APL, but it was checked and kept to a minimum because I had two experts (one a science expert in the field of photochemistry and photobiology, and the other an expert in science education) continually comparing my writing to the primary literature article and providing feedback.

Identification of the Concepts of Evidence Included in the APL

The Concepts of Evidence included in the APL, explicitly and implicitly, were identified. The Concepts of Evidence developed by Gott and Duggan (1995) describe the procedural understanding necessary for working in science disciplines. In 1995, the descriptors could be interpreted as being restrictive in that they were more closely allied to laboratory-based investigations, rather than being applicable to the other types of science-based work. This was especially true where relationships between naturally changing variables were studied, such as in biological fieldwork. More recently, however, Gott et al. (2003) defined the Concepts of Evidence in a way that they could be more readily applied to the range of contexts investigated by biologists and investigators in other branches of science. The compendium of Concepts of Evidence developed in Gott et al. is comprehensive, but as yet tentative, including 21 categories as shown in Appendix G.

I used Gott's categories as a conceptual framework to identify and code the Concepts of Evidence present in the Adapted Primary Literature (APL). According to Gott's classification, the 21 categories of Concepts of Evidence range in complexity from simple Concepts of Evidence such as "observation" and "measurement" to more complex concepts such as “instruments: calibration and error” and “reliability and validity of a single measurement.” Each of these 21 categories is subdivided, and the subcategories are listed under each category as shown in Appendix G. This list of Concepts of Evidence was informed by research and writing in primary and secondary science education, in science-based industry, and in the public understanding of science.

Creation of the Evidence Survey: A research Preactivity

I developed the pre-test and post-test questions of this pre-research activity so that 50% of them examined the substantive knowledge of the APL, while the rest were on the same Concepts of Evidence, directly referred to, or implied, in the APL. The student answers to the pre-test and the post-test together gave me some initial data on how students perceived the Concepts of Evidence and the substantive knowledge contained in the APL.

Crafting the Interview Questions

The interviews that I conducted were semi-structured. I wrote out 11 specific questions for the students and 14 questions for the instructors. However, I remained flexible. If the answers provided by my student participants warranted further clarification or explanation, I asked more probing questions impromptu, thereby making my interviews semi-structured.

Pilot Study

Science 100 (Integrated Science) is an interdisciplinary course offered in the Faculty of Science at the University of Alberta. The pilot study was conducted in the Science 100 class under the supervision of the instructor, who was in charge of this course. Students in Science 100 were given the option of participating in the pilot study or opting out. No students opted out. Over a weekend, students were asked to read the APL and a piece of related secondary literature which were uploaded onto their class’s Moodle site. On the following Monday morning, these students first answered the pre-test, which was distributed to them as hard copies during class time. They were then given a lecture on how this chemistry professor and his group undertook the piece of work described in the APL. Next, they were introduced to Concepts of Evidence in general, and were then specifically taught the ones included in the paper. This was followed by the post-test, which the students completed in class the same day.

I marked the pre-tests and the post-tests immediately. The chemistry professor and I reviewed the questions which were incorrectly answered, checking for ambiguity. Questions which appeared ambiguous or unclear were altered or dropped. We decided to include only the APL in future surveys. The altered pre-test and post-test were then put on-line using SurveyMonkey. I identified the Concepts of Evidence in the APL, and the Chemistry Professor addressed four of these as he taught the teaching intervention. This pedagogic intervention was videotaped, producing a 32-minute-video, and then uploaded on-line. The video was designed as a first-hand account of the major experiments carried out in the APL, as well as an introduction to Concepts of Evidence in general, and a few specific ones in particular.

Site Selection

Having given considerable thought to where I would like to carry out my research project, I chose the province of Alberta where I resided in Canada, and this decision was ratified at my candidacy. I was given names and e-mails of a few faculty members in various higher education institutions in Alberta by the Associate Dean in the Faculty of Science at the University of Alberta. I contacted these faculty members, and I was able to get in touch with 10 instructors teaching Chemistry 101 (taught in the first-year of undergraduate study at the University of Alberta) or its equivalent at eight higher education institutions in Alberta. The fact that the sites of data collection were a mixture of urban and rural institutions, varying in size from small to medium to large, accentuated the diversity of the participants in my study.

Selection of Participants

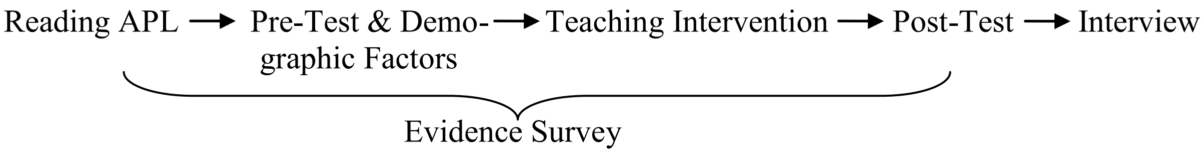
All experiments in educational research share some common features such as offering a treatment, monitoring outcomes of the treatment using assigned units, and thereby comparing the outcomes. In my study, there was at least one treatment (i.e., the teaching intervention), two outcomes that were being measured (i.e. performance on the pre-test and performance on the post-test), units of assignment (i.e., marks obtained on the pre-test and the post-test) and some comparison (i.e., post-test mark minus pre-test mark) from which a change (i.e., Δ_{diff}) was measured (See Appendix H). I took precautions to ensure that all other independent variables and conditions were controlled so that the change in marks (Δ_{diff}) could be attributed to the pedagogical treatment, the teaching intervention.

Random assignment was impossible in my study because the students were given the freedom to choose not to participate, or could opt out of the research at a later stage if they changed their minds. In ideal randomized experiments, however, one is able to infer that the changes observed are caused by the treatment alone because the researchers used random assignment when they initially chose their participants. As Cook and Campbell (1979) claimed, “random assignment is the great ceteris paribus (i.e. all other things being equal) of causal inference” (p. 5). I found it impossible to assign students randomly in my field setting because I was only a guest at the institutions and in the classes where I collected my data. This finding is supported by Cook and Campbell who explained that it is more difficult to assign individuals randomly in field settings than in laboratory settings, because field researchers are often only guests at the sites where they work, while the laboratory researcher, on the other hand, has almost complete control over the setting and acts as the participants' host. Random assignment, therefore, is more frequently achieved in laboratory settings than in field studies, and is more difficult to attain with humans subjects. As random assignment could not be implemented in my study, strategic or purposive sampling was practised in the actual participant selection due to the circumstances under which I had to operate.

Data Sources and Collection

Initially, I gave a short presentation introducing my research project in each of the Chemistry 101 classes of all 10 instructors. These 10 classes together had approximately 300 students. Out of this number, 61 students took the evidence survey. The overall response rate for students, therefore, was 20.33%. The student response rates at the eight postsecondary institutions varied, because there were differences in how the instructors encouraged student participation.

Research sequence. The diagram below demonstrates the research sequence.



The time span from reading the APL to the post-test was a week, usually from one Friday to the next. Because of time restrictions imposed by the instructors, and the limitations to my availability, there was sometimes a two-week gap between the post-test and the interview. The teaching and learning of the content knowledge and of the Concepts of Evidence contained in the APL began at the first stage as the students read it. The learning, however, continued through the pre-test and culminated in the teaching intervention, which clearly identified and addressed the Concepts of Evidence that were included either explicitly or implicitly within the APL. The teaching intervention also covered substantive knowledge contained in the APL in an expository fashion, with diagrams to enhance the understanding of new concepts. The post-test enabled further testing of the students' understanding. The problems students faced, and the erroneous conceptions they harboured, in learning the substantive knowledge of the APL and in their understanding of the Concepts of Evidence were probed and addressed, as far as possible, through the interviews which acted as the final intervention. Each step in the research sequence is described individually in the section below.

Evidence survey: A research Preactivity. The evidence survey is comprised of five items, all of which I created. The Appendix includes four items in the form of the APL (Appendix A), the survey on demographic factors (Appendix C), the pre-test (Appendix D), and the post-test (Appendix D). The fifth item was the teaching intervention which a video sent to the students online as a link. Participating in the evidence survey consisted of the following stages: reading the APL; filling out a preliminary survey on basic demographic factors (such as age, gender, ethnicity, socio-economic status); answering the pre-test; watching the teaching intervention; and, lastly, responding to the post-test. I designed the pre-test and post-test so that one-half of each test was based on the substantive knowledge in the APL, while the other half tested the procedural knowledge embedded in it. The APL treatment was standardized and used in ten classes at eight institutions. At each institution, where data were collected, the APL was uploaded to a website accessible to all students, such as the Moodle site of Chemistry 101 or its equivalent. The pre-test was sent to students as a link to SurveyMonkey, with access limited to just one entry. Once SurveyMonkey alerted me that a student had completed the pre-test, I sent that student the teaching intervention as a video link. The same teaching intervention was sent to all the student participants. The post-test was sent to each student a day later as another link to Survey Monkey. This whole evidence survey was covered in exactly one week at each site, and contributed to my qualitative data.

Teaching Intervention. In my study, students who participated in the evidence survey furthered their understanding of the APL and certain relevant Concepts of Evidence by watching the teaching intervention video. The lead researcher on the project, described in the primary literature by Yarasi et al. (2005), gave the pedagogical intervention. The professor was video-taped as he delivered the teaching intervention in a classroom within 32 minutes. He wrote on the whiteboard as he spoke in a clear, narrative style, describing the whole research activity as though he were telling a story. He interspersed the story with anecdotes which brought in a personal touch into the narration. He initially outlined the specific motivation for that study. It is very hard to detect DNA/RNA damage. The hypothesis they were testing was whether molecular beacons could be used to identify this damage. Thereby, they were hoping to develop new ways to probe damage in nucleic acids.

The teaching intervention brought out the consensually agreed upon model of the molecular beacons, their dark phase and hybrid light phase. The professor drew diagrams representing the hair-pin structure of the dark phase of the molecular beacons combining with the target oligonucleotides to form the opened out hybrid light phase which is fluorescent. These diagrammatic representations showed students how to think of and visualize these changes. This equilibrium which is crucial to understanding the paper was explained in the APL and students would have come to watch the video with their own conceptions of the structure of the molecular beacons. Active learning is brought about in the students by the interaction of their pre-existing knowledge after reading the APL with the new knowledge conveyed through the teaching intervention. There is a dynamic interaction between this prior knowledge from reading the APL and the new knowledge from the teaching intervention, which can bring about conceptual change in the students, in keeping with a constructivist stance on active learning.

Students would have to think for themselves and construct their own understandings of certain specific Concepts of Evidence, which were tested in the pre-test, then explained further in the teaching intervention, tested in the post-test, and finally examined in the interview. During the teaching intervention the professor mentioned the instruments they used in their study such as the absorbance spectrometer, the fluorescence spectrometer, the irradiator and the balance. He talked about the importance of calibrating these instruments and checking whether they had any zero error. In this way he candidly brought in Concept of Evidence #5 which is titled "Instruments: calibration and error," and described how to detect and correct for zero error if it occurred in the afore-mentioned instruments. Students at the first-year undergraduate level are generally familiar with balances and zero error occurring in them, because of their exposure to laboratory work in high school. Again their prior knowledge of and experiences with instruments could interact with the new knowledge in the teaching intervention and consequently promote active learning.

Finally after tracing their main findings in the APL, the professor emphasized the importance of detecting patterns in the data. This neatly highlighted the importance of Concept of Evidence #19 "Patterns and relationships in data." Students are familiar with pattern detection and must have had to practice this concept in their own practical work during their high school years. Again the students' prior understanding would interact with what was taught during the teaching intervention and promote active learning. By participating in my whole research activity therefore students would have been compelled to construct their own understandings of these particular Concepts of Evidence which would be consistent with what I understand to be a constructivist approach.

Instrumentation: Individual Semi-structured Interviewing. Individual, one-on-one interviews are commonly used as a data source in qualitative research because the descriptions they provide of contexts, themes, and issues are generally rich and holistic (Merriam, 1998; Creswell, 2008). Before visiting a postsecondary institution where I was scheduled to interview for one or two days, I contacted each participant at that site and arranged a time that was mutually convenient. I estimated each interview to take about 40 to 45 minutes. In order to preserve consistency, I made a number of interview questions as described earlier under "Initial Preparation" (see Appendix E), which were piloted using participants from Science 100. From the initial answers, I asked leading questions, or provided prompts, which would probe further the students' understanding of the content of the APL or their understanding of the Concepts of Evidence. This enabled the flexibility of semi-structured interviewing. In almost every institution, the instructors that I contacted provided me with a quiet room where I was able to conduct my interviews. This was crucial because the interviews were audio-recorded. In the case of a couple of students, the interviews were conducted over the phone using the speaker phone mode. On completing the interview, each student participant was given a gift certificate as a token of appreciation and thanks.

I also prepared interview questions to ask the instructors (included as Appendix E) in each of the eight institutions, and piloted them as well. I interviewed the instructors in the same room as the students or, in some cases, in their offices, and I also audio-taped these interviews. Then I transcribed the 55 student and 9 instructor interview audio-recordings and notes. While transcribing, I noted my initial impressions after listening to each of the interviews, in a researcher's journal. I listened to the interviews a number of times, in order to immerse myself in the data. I also turned to the literature to see if my findings could be supported.

Analysis of Data

In my study, data collection and data analysis occurred almost simultaneously as recommended by Creswell (2008). Once each transcription was completed, I noted my initial impressions of the interview in my researcher's journal. There were two parts to the data analysis – the qualitative part from the transcriptions of the interviews and the qualitative analysis of the evidence survey. The transcriptions of the interviews were treated as the primary data in this study.

Thematic Content Analysis for Qualitative Data

The approach that I used in my study to analyze my qualitative data was thematic content analysis which measures psychological characteristics of individuals, usually for experimental, quasi-experimental or *ex post facto* research purposes. Thematic content analysis is a type of psychometric analysis with inbuilt measures intended to reveal internal, psychological constructs. Content analysis itself is a research technique for making replicable and valid inferences from texts to the contexts in which they are used (Krippendorf, 2004). One of the objectives of content analysis is to trace the psychological characteristics of the groups or the individuals who were responsible for creating these texts (Neuendorf, 2002). In my study, the transcripts from the interviews were the participants' own messages and these replace other possible measures such as self-reported scales and indexes. The contemporary source of reference for this technique is from the book, *Motivation and Personality: Handbook of Thematic Content Analysis* (Smith, 1992). According to Smith, the purpose of thematic content analysis is to score messages for content, or for style, or for both. "It is done for the sole purpose of assessing the characteristics or experiences of persons, groups, or historical periods" (p. 1). In my study, thematic content analysis was employed to fathom the extent of the learning of

Concepts of Evidence achieved by the participants, and the experiences that these students and instructors went through as they partook in the evidence survey and the interviews. I broke down the transcription or "communication" into bits that could be counted, a process known as "unitizing" (Carney, 1971, p. 52). Although the counting units may vary, and may be words, characters, themes or "interaction-units," in my study I counted themes (p. 52).

Analysis of the Evidence Survey

The first step in the qualitative data analysis of the data obtained from the evidence survey was to print the pre- and post-test responses from Survey Monkey. I devised a marking scheme for both tests and had the same total mark assigned to each. I then marked the pre- and post-tests of all 61 participants, based on this marking scheme. In order to answer my second research question, I had to probe the initial understanding that first year undergraduate students have of the Concepts of Evidence and the content knowledge found in the APL. This is an onerous task which to my knowledge has not been investigated as yet. I had to start somewhere and I decided to set the pre-test as my baseline. Since the students took the pre-test before the teaching intervention, their performance on this test seemed an appropriate measure of their inherent ability to read and understand a science article and make sense of it on their own. This would shed light on their individual learning styles and how each student best learned the content as well. The post-test, on the other hand, was taken after the pedagogical intervention. It would, therefore, reveal how the instructors' teaching methods contributed to the post-test scores. There could also be other factors contributing to the student performance differences in the post-test such as: the help that students got from their peers; the guidance offered by their instructors during the evidence survey; the notes that students had taken during their chemistry, genetics or biotechnology courses; and the online searching that students could have done to help them understand any part of the evidence survey. By adopting such an approach the analysis of the evidence survey was what I understand to be constructivist.

Researcher's Predictions

Before I engaged in data collection, I had hoped to find that Concepts of Evidence could be both taught and learnt, and that pedagogy had a significant impact on the actual extent of the learning that took place. This is because, from what I have experienced in high school and as a university student in India, in the United Kingdom and in Canada, I believe that good teaching always enhances student-understanding. The teaching intervention, I thought, would improve the students' learning of the Concepts of Evidence, and would display a strong interrelationship with the APL. From my learning experiences in different parts of the world, I expected the majority of students to find the teaching intervention very useful, because of its clear, succinct explanations and illustrative diagrammatic representations. For the biotechnological methods used, diagrams are ideal for conveying changes which cannot be observed. I thought the teaching intervention, would be effective in bringing about conceptual change in the learning of substantive knowledge and the Concepts of Evidence, included in the APL.

I expected that most students would not have a strong opinion about Concepts of Evidence being introduced as part of their curriculum in first year courses, because they would not have really experienced this in their own lives. I expected most instructors, on the other hand, to support introducing Concepts of Evidence in their teaching of chemistry at the first-year undergraduate level. Qualitative methods (i.e., semi-structured interviews and the evidence survey) were used to test the students' understanding of Concepts of Evidence in my project. I expected the information gathered from these methods to complement each other.

I thought most students would struggle through the APL, but those who persisted would find it rewarding, because of the scaffolding provided by the APL. They would find the APL daunting at first, and do quite poorly on the pre-test. I expected different sections of the APL, to evoke different challenges. The biotechnological methods used in the APL, for example, were demanding as they built on a certain level of prior knowledge in genetics. Based on constructivist principles, however, I would expect most students to find the teaching intervention very educative and effective. Many students I think may find the post-test easier, and may perform better on it, based on the extent of the interaction between the APL and the teaching intervention. Since all the students were given the same teaching intervention, the degree of interrelationship between the APL and the teaching intervention would determine the extent of the learning that occurred. The greater the dynamic interrelationship between the APL and the teaching intervention, the wider would be the extent of student learning, and the deeper would be the scope of student understanding. I expected students to clarify some of their initial erroneous conceptions through the interviews, because I believe that one-on-one teaching is the most effective pedagogical tool for addressing and correcting misconceptions. Following an examination of how the data were analysed, it is important to establish how valid and reliable the data collected in this qualitative study are. The section below is dedicated to this task.

Data Quality Issues

According to Freeman, de Marrais, Preissle, Roulston, and Sr. Pierre (2007), all data are either constructions or interpretations that are produced during social interactions. Consequently, “pure” and “raw” data which are “uncontaminated by human thoughts and actions” just do not exist (p. 212). My qualitative data, therefore, are interpretations made by participants as they respond to my questions or, by me as the researcher who notes their observations. “Neither research participants nor researchers can be neutral, because ... they are always positioned culturally, historically, and theoretically" (p. 212). In qualitative studies, the participants' interpretations are considered data. It is vital, therefore, that I as the researcher "capture these interpretations accurately and without distortion" (Teddlie & Tashakkori, 2009, p. 212), interpreting the data to a minimal extent. This can only be achieved only if researchers maintain carefully the trustworthiness criteria outlined below.

Trustworthiness criteria.

Similar to Marshall and Rossman (2006), I was convinced that the value of my research resides in the trustworthiness of its findings. Trustworthiness is a global qualitative concept introduced by Lincoln and Guba in 1985. According to them, four questions can be posed which formulate the criteria for establishing the "truth value" (p. 290) of my study. First, how can the credibility in the "truth" of the findings of my study be established? Second, how applicable are my findings in another context, or how transferable are they among a different group of people? Third, how can I be sure that the findings would be replicated if this study was carried out again, with the same participants in the same context? Fourth, how can I be sure that the findings of the study are a true reflection of the participants' views, and not contrived by my biases, my interests, my motivation or my perspectives as the researcher? Each of the four questions can be summarized in order by one of the four criteria - credibility, transferability, dependability, and confirmability. Trustworthiness envelops all four of these criteria, and is defined as "the extent to which an inquirer can persuade audiences that his or her findings are 'worth paying attention to'" (Teddlie & Tashakkori, 2009, p. 296; Lincoln & Guba 1985, p. 290). What these four criteria imply and how I have adopted them in my study are considered next.

Credibility. In order to maintain the "truth value" in my study, I had to demonstrate that the reconstructions that I had arrived at as a result of the process of inquiry are credible to my participants who are the constructors of these original multiple realities (Lincoln & Guba 1985). In other words, I had to ensure that I, as the researcher, had "represented these multiple constructions adequately" (p. 296). The three techniques that I implemented in my study to enhance the credibility of my findings were: (a) prolonged engagement, (b) triangulation techniques, and (c) member checks.

Prolonged engagement. It is vitally important that researchers spend a sufficient amount of time in the field with the intention of building trust, learning the "culture," and testing for misinformation which they may have gathered either from informants or from the biases that they themselves possess. The reason for having prolonged engagement is to widen the scope for researchers by "making them aware of the multiple contextual factors and multiple perspectives of participants in any given social scene" (Teddlie & Tashakkori, 2009, p. 212). In my study, I first visited the sites of data collection and made contact with the heads of the Departments of Chemistry. I briefly introduced my research project to each Head and then proceeded to meet with the actual instructor responsible for teaching the Chemistry 101 course or its equivalent. I had much lengthier introductions with the instructors as I explained my project in detail to each of them, and came to agree on some time frame to introduce my project to the students in their respective classes. On another pre-arranged day, I visited each of the Chemistry 101 classes and explained the role of evidence in science. I then introduced my study to the students explaining what participation in it entailed for them. I did this initial introduction presentation with every class from which I enlisted participants for my study. The preliminary introduction also enabled me to generally observe the classes from which my participants were recruited. I found that the students in all the classes were very receptive to my presentation, and listened intently with interest. However, the response rates of the students were low and one of the main factors raised by my participants was the lack of time as some were having very busy schedules with their assigned coursework. In addition, I observed that the response rates of the students were much higher in the rural postsecondary institutions than in the urban ones. This could be because the instructors in these rural institutions promoted this study and encouraged participation by donating some marks when their students successfully completed my project. In addition, I found that the students in one rural institution were remarkably keen and punctual in participating in all stages of my research project.

After the introduction, I made the APL available online to all students, enlisting those who signed the consent form. I then administered the pre-test, teaching intervention, and post-test online to everyone who enlisted. I visited each site a third time to interview the participants and their instructors. Through all these stages I was having a prolonged engagement with each of the student participants and their instructors in a variety of ways: face-to-face, via e-mail, and using SurveyMonkey. I interviewed three students over the phone, using speaker phone mode, because they were absent when their face-to-face interviews were actually scheduled at their institutions. I was able to observe each student's behaviour online, when responding to the pre-test and post-test, and continued to interact with each of them face-to-face, during the interviews at their respective institutions.

Triangulation techniques. Triangulation techniques are used when multiple data sources, data collection procedures, data analysis steps, research methods, investigators and inferences are combined and compared at the end of a study. It was Webb, Campbell, Schwartz and Sechrest (1966) who made the first reference to triangulation, stating:

> Once a proposition has been confirmed by two or more independent measurement processes, the uncertainty of its interpretation is greatly reduced. The most persuasive evidence comes through a triangulation of measurement processes. (p. 3)

In the 1970s, however, Denzin (1978) advanced the discussion on triangulation to include four different and distinct types: (a) data triangulation**,** (b) investigator triangulation, (c) theory triangulation, and (d) methodological triangulation. Data triangulation was implemented in my study because the single problem of fathoming the understanding of the Concepts of Evidence held by first-year undergraduate students was explored by the qualitative methods of interviewing and conducting a survey (i.e., the evidence survey consisting of a pre-test and a post-test). Consequently a single research question or problem was studied using two separate, qualitative data collection methods. Multiple sites (i.e., eight different postsecondary institutions across the province of Alberta, Canada) were also used to carry out the data collection, because I as the researcher looked for "as many different data sources as possible which bear upon the events under analysis" (Denzin, 1978, p. 301). By triangulating data sources in this fashion, I could "efficiently employ the same methods to maximum theoretical advantage" (p. 301). Furthermore, by selecting these dissimilar settings systematically, I was able to discover what the problem I was investigating (i.e., the understanding of Concepts of Evidence held by first-year undergraduate students) had in common across these settings.

In my study, I aimed to seek multiple realities rather than trying to discover "one truth," "one reality," or a "final solution" (Peshkin, 1993, p. 28). It must be borne in mind that several qualitative researchers do not believe in a "single reality that can be triangulated" (Teddlie & Tashakkori, 2009, p. 213). I interpreted the differences in the way events and phenomena are represented, "as the alternative realities of the participants" who are included in my study (p. 213).

Member checks or Inter-rater coding. For my study, this strategy was adopted for ensuring data quality. It involved asking colleagues in the same social environment to substantiate my representations of events, behaviours, or phenomena. This is the most important strategy (Krippendorf, 2011) for establishing the credibility of the interpretations that I as the researcher have made of the perceptions of my participants. I asked one of my colleagues, who was my senior and who had completed his PhD recently, to do a member check using my data. I provided him with my interview questions, the students' transcripts, my codebook, and some sample coding forms. My colleague then proceeded to code a few of the students' transcripts, and compared his coding to mine. There were no marked differences in the coding. This type of informal member checking is one of the most crucial techniques for establishing credibility.

Transferability. Transferability is the criterion that governs the "transferring of inferences from a specific sending context to a specific receiving context" (Teddlie & Tashakkori, 2009, p. 296). Lincoln and Guba (1985) claimed that qualitative data could "only set out working hypotheses together with a description of the time and context in which they were found to hold" (p. 316). I found that this claim applied to my qualitative data as well. However, the question of whether these working hypotheses will hold in some other context, or even in the same context, but at some other time, is an empirical problem. The resolution of this empirical issue will depend upon "the degree of similarity between the sending and receiving contexts" (p. 316). This degree of similarity can only be gauged by thick description, a technique that I implemented in my study for enhancing the transferability of my study.

Thick descriptions. I used this strategy, which was emphasized by Teddlie and Tashakkori (2009), in order to provide evidence about "the transferability of interpretations and conclusions from qualitative investigations" (p. 213). Lincoln and Guba (1985) maintained that it "is the responsibility of the inquirer, to provide a person contemplating application in another receiving setting, to make the needed comparisons of similarity" (p. 359). In addition Teddlie and Tashakkori (2009), mentioned the importance of making detailed descriptions of “the context and other aspects of the research setting” (p. 296). I implemented this suggestion by recording such necessary detailed descriptions in my research journal, so that other researchers could make the necessary comparisons within the contexts in which they are working.

Dependability. Dependability addresses “the extent to which the process of the inquiry is dependable” (Teddlie & Tashakkori 2009, p. 296). It is a qualitative analogue for the quantitative concept of data reliability. Dependability is a measure of the extent to which variations that arise in a phenomenon can be accounted for consistently by the “human instrument” when placed in different contexts (Teddlie & Tashakkori, 2009, p. 209). Since I made the assumption that I was measuring what I intended to in my study, the question I addressed while conducting my data collection was whether my measurement or recording was consistent and accurate, in keeping with the conditions offered by Teddlie and Tashakkori. If the results are dependable, then variations across different qualitative contexts will be tracked consistently.

Dependability audit. I employed this technique to enhance dependability in my qualitative data. It probes into my process of inquiry, examining whether the “inquiry decisions” and the “methodological shifts” (p. 295) that I made were appropriate. I made journal entries of my interviews, and always checked whether the decisions I made while interviewing, and the methodological shifts I made, were appropriate: I usually found them to be so. If I noticed any inappropriate methodological shifts, I noted them and avoided them in future interviews.

After transcribing, I read through my transcripts intently and studied them. Since I had planned and created the interview questions, I was able to follow each interview carefully, and spot the leading and probing questions, which exposed the understanding and learning of students. The decisions to ask these types of questions in a timely fashion were the “inquiry decisions" that I had made. I always checked the slight or major "methodological shifts" made because of these questions, and found that these shifts brought the learning back on track, enabling me to continue with the interview questions designed.

Confirmability. Confirmability is indicative of how confirmable the products of the inquiry are. I looked into aspects such as whether the results I obtained were backed by sufficient data, whether the inferences made were logical, and whether the results obtained were tainted by my bias, and so on. Techniques such as a confirmability audit, triangulation, and keeping of a reflexive journal are used to enhance confirmability. Out of these, the techniques that I used in my study were triangulation and reflexive journalling.

Reflexive journalling. This strategy gives information about all the four criteria for trustworthiness mentioned earlier. According to Lincoln and Guba (1985) the reflexive journal is "a kind of diary in which the investigator on a daily basis, or as needed, records a variety of information about *self* ... and *method"* (p. 327, italics in original). As far as self is considered, the reflexive journal can be envisaged as offering the same kind of data about the human instrument as is often obtained from the paper-and-pencil or other instruments employed in conventional studies. When considering method, the journal is a trove of information about the methodological decisions made and the reasons for making them. Such information would be of significant value for the auditor. (p. 327). In my study, I wrote down in a reflexive journal the initial impressions I had of each of the interviews and of the participants immediately after interviewing them. I also wrote down any methodological decisions I made, particularly noting why I made them. I noted that I was quick to gauge whether the interview would proceed in a fruitful direction as expected. I was able to monitor the situation effectively, and decide what had to be asked next. I was quick to ask probing questions, and get the discussion back on track.

Many of these probing questions clearly revealed the extent of the students' understanding. I noted that I was very comfortable with the subject area covered. Occasionally, I felt that the students had not understood the experiments outlined in the APL. However, no matter how they came, I always tried to promote their learning of the Concepts of Evidence. I enjoyed the interviews and found some of them very rewarding and worthwhile. I noted that I had to be very careful and consistent with terminology in order to avoid confusion and enhance clarity (e.g., when dealing with the question on sampling during the interviews). Despite the efforts outlined above to increase its trustworthiness criteria, this study does have its limitations and delimitations as outlined below.

Researcher's Bias

As the researcher, especially in the qualitative research paradigm, one cannot separate oneself from the topic one is studying. It is actually in the interaction that occurs between the researcher and the researched that knowledge is created (Mehra, 2002). In the interviews, it is actually by asking the question, and getting an answer from the participant that the argumentation starts. It is in the nature and progress of this argumentation that knowledge is created about the understanding that the participant has about certain Concepts of Evidence. Therefore, "the researcher bias enters into the picture even if the researcher tries to stay out of it" (p. 9).

One of the major objectives of any scientific investigation is to describe or explain a phenomenon in such a way that it avoids the biases of the researcher (Neuendorf, 2002). In other words, objectivity is desired. Phillips and Burbules (2000) explained this stating:

> The activities of science should be internally directed by the values – the so-called cognitive values – that are relevant to the aim of science to produce warranted knowledge; these activities should not be subverted by external and irrelevant values and biases. (p. 56)

I have tried to confront and be wary of my biases.

However, as Berger and Luckman (1996) pointed out, there is no such thing as true objectivity in research. “Knowledge and facts are what are socially agreed on” (Neuendorf, 2002, p. 11). From this viewpoint, all acts of human inquiry are innately subjective, but we have to make a concerted, strenuous effort for consistency among our inquiries. “We do not ask ‘is it true?’ but rather ‘do we agree it is true?’” (p. 11). Researchers and scholars referred to this standard of practice as *intersubjectivity* (Lindlof, 1995). I have adhered to this standard of intersubjectivity in my study by implementing intercoder or interrater reliability tests, which will be described later in this chapter.

As I was very aware of the power I had as the researcher in the dynamics of the researcher-student relationship, I attempted to be as neutral as possible while conducting my interviews, a tactic adopted by Mehra (2002) and others. I made an extra effort to get the participants to recount their perspectives without requiring any approval or confirmation from me. I was especially careful to ask semi-structured questions, being cautious not to steer the participants in a direction that may have seemed to endorse a particular answer or response. I tried my best to model my interviews after the conversations that would take place between two parties that trust each other. Behaving as trusting parties required that I be up front and honest about my bias when discussing research design and particular methods used in the APL.

Science is additionally protected from being intruded by “nonepistemically relevant values” and biases by the fact that this subject “is organized as a communal activity,” (Phillips & Burbules, 2000, p. 60) having a strong tradition of open inquiry and discussion, of replication, of peer review and so on. It is this openness to criticism promoted by science that acts as the best safeguard against errors, assumptions, values, and biases, because these will get rooted out and exposed to the light during scientific discussions. As Popper (1945) emphasized:

> It may be said that what we call "scientific objectivity" is not a product of the individual scientist's impartiality, but a product of the social or public character of scientific method; and the individual scientist's impartiality is, so far as it exists, not the source but rather the result of this socially or institutionally organized objectivity of science. (p. 374)

Since bias cannot be justified by evidence or argument, it is important to limit it, and keep it to a minimum. One can, therefore, define accuracy as the extent to which a measuring procedure is free of bias.

I took great effort to minimize nonrandom error in my qualitative data collection and interpretation. I listened to the tapes of my interviews and read the transcripts from these tapes several times, thereby immersing myself in the data. I tried to keep an open mind in listening to the transcripts of the interviews, considering that in any problem or situation there is not just one truth, but many truths (Phillips & Burbules, 2000). On exposing myself repeatedly to the data, I listened carefully and asked myself whether I had missed anything in previous readings and listening. I made an extra effort to report on the full spectrum of answers to each question, ensuring that all responses were included. I even enumerated the number of answers that fell into each category that constituted the spectrum.

Harding (1993) expressed the concern that there are deep and hidden sources of bias, which are unlikely to be identified, and eradicated if we rely only on the method of open criticism vouched by Popper (1945). As cited in Phillips and Burbules (2000), Harding argued, "there are deep male biases built into the very category systems used in science" (p. 61). These biases, however, are taken for granted and accepted within the predominantly male scientific community. Consequently, these biases are not likely to be challenged or removed, simply because they are not perceived as biases at all, but have become a part of the "unquestioned cultural background against which contemporary science is practised" (p. 61). Harding (1993) recommended what she calls “strong objectivity” (p. 237) which can grapple with these deep-rooted sexist and other biases. The way to achieve this “strong objectivity” according to her is “to enfranchise the marginalized group” (p. 237) by incorporating them as full members of the community. It is only the marginalized people, at the bottom of the social ladder, who can see certain things as problematic. For example, certain values held by the dominant group have become so incorporated, that they are not noticed by people at the top of the social ladder. In my research project, minorities are well-represented among my participants. The views and learning displayed by minority groups are therefore, well-documented in my study, thereby reducing bias. The bias that I could not avoid, however, was that brought in by the theoretical framework of constructivism on which my study is based. The first tenet of this theory is that each learner constructs an individual understanding mentally, and does not just receive knowledge passively (Driver, 1981; Fensham, 1992a). It is vital, therefore, to further examine the limitations and delimitations of this study.

Limitations of the Study

In order to find out whether conceptual changes in the understanding of Concepts of Evidence can be reversed, it would have been best to monitor this understanding using a longitudinal study covering five or seven years. Since this is only a one-time, study, I cannot really monitor the cognitive development of the students in their understanding of Concepts of Evidence across years.

In my research project, Concepts of Evidence are introduced and interpreted to the students through their teachers. The teachers and instructors bring their own natures, beliefs, and value systems and these will affect the teaching and the learning of Concepts of Evidence. The students’ understanding of these Concepts of Evidence and their attitude towards them will, therefore, be influenced by their teachers’ values, beliefs and attitudes towards these Concepts of

Evidence. From my own personal experience as a teacher, I know that I bring to my classroom certain values that I treasure, such as integrity, justice, and fair play, especially in matters associated with teaching and learning. I try to communicate to my students my belief that they are unique and valuable in their journey to understand the concepts in science, and especially in chemistry. Consequently, some inevitable bias is automatically introduced into students' understanding of Concepts of Evidence. Some teachers may highly value procedural knowledge and may consider the teaching and testing of students' understanding of Concepts of Evidence as essential and vital, while others may see it only as one option, among many others, for improving the pedagogy of science. This is an inevitable bias that will be present in the study.

Delimitations of the Study

One of the delimitations of this study is that I have chosen only first-year undergraduates who are taking science courses, particularly Chemistry 101 or its equivalent. The reason for this delimitation is that since the curricula of the first-year undergraduate chemistry programs have placed a greater emphasis on procedural knowledge, it will be easier to accommodate a teaching and testing of the understanding of the Concepts of Evidence by aligning them to suitable points in the curriculum.

A second delimitation that this study abides by is that the samples are observed only in chemistry classes, and not in any of the other science disciplines. This is because the study is looking specifically at the understanding of Concepts of Evidence in chemistry held by university undergraduate students taking Chemistry 101 or its equivalent.

A third delimitation that applies to this study is that it is piloted in a Science 100 class. Since this is an interdisciplinary science course introduced in 2009, and offered at the University of Alberta, students taking this course have all passed Chemistry-30 (Grade 12, Chemistry), and are accustomed to doing questionnaires in various disciplines of science. Therefore, they will also be adept at answering the questionnaire used as the data collection instrument in this research project. This Science 100 class was, therefore, an excellent venue to develop my questionnaire.

A fourth delimitation is that there are four different types of knowledge. When justified, knowledge turns into evidence. Although it was important to mention all the different types, only two – substantive knowledge and procedural knowledge – are directly relevant for this study. Strategic and situational knowledge are essential for higher order cognition in science. Yet because they are not the basis of this project, they are not discussed at length in the literature review in Chapter Two.

Summary

The methodology that I have used in my study is based on constructivism. This theoretical framework has been instrumental in designing the methods that I utilized as well. I have traced how I entered the study, the research design employed and the methods of data collection, data analysis and data interpretation deployed. I analysed the qualitative data that I collected from the interviews and the evidence survey using thematic content analysis. Much effort has been taken to put in place measures to ensure that my study can be held up over time to be trustworthy. Attempts have been made to minimize my bias, to predict the findings, and to be very cautious with the extent of generalization because my study is exploratory in nature. Despite all the precautions taken, this project has its limitations and delimitations, which must be considered when analyzing and interpreting the results. Some of the limitations and delimitations need to be addressed with further research.

CHAPTER FOUR

FINDINGS OF THE STUDY

Prelude: The Examining

Slowly, the tall sheaves of the rice plants ripen hanging in golden bunches. The green plants with pendulous inflorescences have turned into yellow plants with husks. When the husks ripen, the paddy fields turn a golden yellow. Just as the mature husks have to be examined, in order to reveal the precious rice grains housed within, the products of data collection in a research project have to be analyzed to reveal the findings of the study.

.

This chapter conveys the findings of my data collection. The presentation of findings is divided into two sections: the student interviews and the interviews of the instructors. The data, resulting from the semi-structured interviews of the students as well as the instructors, were analyzed using the technique of thematic content analysis, which is described below. The themes that emerged from the content analysis are discussed. The pre-tests and post-tests which were marked and analyzed also contributed to the qualitative data.

Thematic Content Analysis of Data

The data for the thematic content analysis were obtained by conducting semi-structured interviews with 55 students and 9 instructors. I audio-recorded the interviews and then transcribed them. I then coded the transcriptions from the students' interviews onto Students Coding Forms using a Students' Codebook. Similarly, I coded the transcriptions from the instructors' interviews onto Instructors' Coding Forms using a separate Instructors' Codebook. I created both sets of coding forms and codebooks (See Appendix F).

The Codebook was constructed by listing the interview questions and their sub-sections in chronological order. Under each question, all the possible responses were listed as multiple choice answers, which were numbered. The Coding Form translated each answer to the interview questions to a numerical code from 1 to 10. Codes 1 to 8 were the various, possible answers to questions, leaving some leeway to accommodate any unexpected or undeterminable responses. The answers each student provided to the questions were coded and entered into the form with a tick against the appropriate answer(s). If there were any extra points to note about the answers, they were put down at the bottom of the Coding Form. In addition, the Coding Form also provided the page number in brackets, where the answer to each interview question occurred within the transcripts.

As the transcripts were analyzed, the responses to questions were underlined in red, sectioned off, and numbered with the sub-sections clearly indicated, so that it was easy to trace the answer to each and every question. The sub-section to which each answer belonged was also made easily apparent. In some of the interviews, interesting points were raised outside of the actual answers to the interview questions. These areas in the transcription were sectioned off and then copied to create a bank of verbatim quotes. These verbatim quotes were not part of the immediate answers to the questions, but were developed from them by appropriate questioning on my part, as the interviewer. The interesting ideas that arose from these questions and answers have been included in this chapter.

Analysis of Student Interviews

After coding the student interviews, an overall representation of the students' responses was attained by filling out Analysis Forms for each question, based on the information gathered in the Coding Forms (Neuendorf, 2002). The Analysis Forms, which I made and filled out (See Appendix I), were further analysed to identify the main findings that arose from the interviews. Those findings, which clustered together, sharing a common and recurring idea, were recognized as a theme. A theme was, therefore, the central, common notion that emerged from each group of findings. The four main themes which emerged from the students' interviews and the five themes from the instructors' interviews are discussed in detail below, along with the group of findings which contributed to each theme. I have arranged the themes sequentially so that Theme One is in the response to Research Question One; Theme Two is the answer to Research Question Two and so on up to Theme Four. In the instructor interviews, Theme Five moved beyond the research questions and included the suggestions offered to improve this research activity.

Once each finding was established, its frequency of occurrence was determined, and in this way the data were *quantitized* (Miles & Huberman, 1994). This was done for each of the findings obtained in this study. This is referred to as Conversion Mixed Data analysis (Sandelowski, 2001). Here, collected qualitative data are converted into numbers (i.e., quantitized). This prevents the researcher from "overweighting" or "underweighting" (p. 234) the emergent findings.

Theme one: Efficacy of the teaching intervention. The second theme that emerged is given the heading "Efficacy of the teaching intervention" and is followed by a general description of what this theme represents. Ensuing this, each of the findings which contributed to this theme is examined more closely. This first theme deals with one of the most important features of this research - the teaching intervention - which is the most direct pedagogical assistance given to the students with the specific goal of helping them in their endeavour to comprehend the APL and develop an understanding of the Concepts of Evidence entrenched in this paper. The pursuit of this goal was examined from various angles such as the overall impression the students had of the teaching intervention, their evaluation of its usefulness, and as assessment of how they perceived their performance on the post-test. These three overarching perspectives were the three findings discussed below.

The first finding: Overall impressions of the teaching intervention. The students were asked their opinions about the teaching intervention. Over four fifths of the students (See Appendix I) found it useful, with more than half readily referring to it as "excellent" or "very useful." According to Student 5, "Yes it (i.e., the teaching intervention) made a big difference. That's when I understood" (p. 17). The teaching intervention played a substantial role in furthering the students' understanding of substantive knowledge and also about the Concepts of

Evidence. This sentiment was echoed by Student 9, who claimed that out of the whole research activity, “I liked the video best” (p. 29). Student 16 found the teaching intervention very useful because it was much more in depth than the APL. Commenting on the teaching intervention Student 18 said, “... It wasn’t too long. I mean it was something I wasn’t extremely interested in but it took a lot of effort to actually you know, watch it for that long” (p. 21). The length of the video seemed to be an issue here.

Unreservedly, Student 20 remarked, “It totally was helpful after reading the article, and everything, like, finding out these Concepts of Evidence are the things he was looking for” (p. 39). Student 23 commented:

> It definitely helped to find out all the actual definitions of the Concepts of Evidence, like clarifying what an independent variable is, a dependent variable is, figuring out what defines a fair test, ... the only thing I didn’t like about the video was that when the camera was moving, it kept on clicking. (p. 26)

This student appreciated the clarification of the basic concepts such as dependent and independent variables and the fair test.

When asked whether the whole research activity was a good experience, Student 31 replied, “It definitely helped, yeah. I watched the video and understood most of it ...” (p. 52). The teaching intervention, therefore, had a definite role in promoting the understanding of these students. Student 32 found the video very well organized. An interesting comment was made by Student 52 when asked whether the teaching intervention was useful. This particular student replied, “I’d say so. I kind of learned a few things about myself” (p. 39). This is a very revealing statement, shedding light on the actual experience of going through the teaching intervention. While participating in the research activity, Student 52 not only learnt the substantive knowledge and the knowledge of the Concepts of Evidence, but also learnt about self. This student does not say what exactly was learnt, but it certainly does pique one’s curiosity. Student 52 must have found the discovery about self very interesting, because this student remembers it.

***The second finding: The teaching intervention as a pedagogical tool*.** The students made a range of comments on the pedagogical usefulness of the teaching intervention, which addresses research question two: these comments comprised the second finding. Roughly two-fifths of the students greatly appreciated and commended the drawings and flow-charts presented on the board during the teaching intervention because they found them to be both useful and informative. For example, Student 30 claimed:

> The video was the most positive aspect in that it relayed and provided auditory teaching. It provided visual imagery. And so, rather than just sitting and reading, it provided different stimuli, which I feel increases the retention and ability to understand. I guess, it's like someone breaking it down and explaining it. ... Certainly maybe it is repetitious, because they are presenting the same data, but it allows for the mind to process it, in a different way, in different pathways. Ears versus the eyes, it is very, very useful. (p. 43)

The auditory and visual stimulation achieved by using the video in teaching about the content knowledge and the Concepts of Evidence was found to be effective by Student 30. Processing the same information in various ways using different faculties, including the faculties of sight and hearing, can profoundly enhance the understanding and retention of what is presented. For many students, the difference made by the teaching intervention was so remarkable that "It was like the light came on" (Student 26, p. 9), like the difference between night and day. "I thought it was excellent," remarked Student 40 (p. 42).

When asked initially about the teaching intervention, Student 56 had quipped "I really enjoyed it, actually," and went on to substantiate how it was useful, saying:

> S56: I just think it was a lot easier to understand, when you are actually focussed on a person that is describing it to you. He drew everything on the board and made a lot more sense, and was easier to follow along ....
> [1]EV: That was helpful for you?
> S56: Yes. (p. 29)

---

[1] EV are the initials for myself, Elizabeth Vergis

This student appreciated the fact that the teaching intervention was delivered in a teacher-centred fashion.

Almost three-quarters of the students felt that the most beneficial part of this research activity was the video, and gave varied reasons for this choice such as: the video was easy to follow; it covered the background information very clearly; it summarized the main points; it was very well-organized and, the information contained in it was very useful. If there were errors in the paper, they were explained in the video. The pace of teaching in the video was not rushed and was ideal to just sit and watch (Student 44). It was "a nice change of pace from having to just read all the time. You can watch the lectures" at your own convenient time (Student 53, p. 50), in your own space. "It explained pretty much step-by-step what was going on" (Student 46, p. 34). One student felt that the "conversational style" of the video made it easier to understand, especially "when they use diagrams and, like, use arrows on it, this is what this means and this is how we did it, and stuff like that" (Student 49, p. 11). Student 30 summed it up best, saying "The video was awesome" (p. 30).

One aspect of the teaching intervention that Student 20 appreciated was that "it pointed out what we were supposed to be looking for in the article" (p. 39). Another student (Student 23) was sure that "It definitely helped to find out all the definitions of the Concepts of Evidence" (p. 26). When asked what was most helpful about the teaching intervention, Student 43 replied:

> Like, the pictures and stuff, and he would describe what would happen when the thing joins with the molecular beacons. They try to describe it in there but it is very wordy. So he just drew it,showing how it worked and then describing the graphs and how they worked as well. (p. 8)

This student found the APL description about the molecular beacon binding with DNA or RNA to be wordy and confusing. When this was drawn on the board during the teaching intervention, the explanation of what happens became lucid. "It helped me understand, you know, put pictures to things that I might not have thought visually in the head from reading the article..." said Student 51 (p. 34).

The third finding: Performance on the post-test: Student perceptions. Almost three quarters of the 55 students thought they had fared better on the post-test, and attributed this to the teaching intervention. Some, such as Student 51, remarked "I think the teaching intervention helped to a point" (p. 34). The majority of student participants thought that the teaching intervention enhanced their understanding of both the substantive knowledge and the Concepts of Evidence covered in the APL. This, in turn, according to the students, increased their confidence about taking the post-test. About a quarter of the students were of the opinion that their confidence level was boosted by the teaching intervention and also by the fact that they had done the pre-test: this heightened confidence and enabled them to perform better in the post-test. Most of the students who participated were of the opinion that the teaching intervention was a remedy of great efficacy: that was a great compliment!

The pre-tests and the post-tests were marked, but the students were not shown the marks that they received. When they were asked about the post-test, almost three fifths of the students said they thought they had performed better on it. Four students out of 55 confessed that they felt that they had not done better on the post-test. "I found the pre-test a lot easier," said Student 8 (p. 24). Quite a few students expressed uncertainty, but thought they probably had done better on the post-test. "I don't know. I think so," was the answer offered by Student 17 (p. 13), while Student 16 said, "Maybe, I am not really sure" (p. 5). One participant expressed the opinion "I don't know if the test results would be different" (Student 21, p. 6).

There were those who had a different view, such as Student 5, who stated, "The pre-test had harder questions than the post-test" (p. 17). The same sentiment was expressed by Student

41, who said, “I found it (post-test) a little more difficult, but I felt they were similar enough ... it was still easy” (p. 5). Another student reiterated “For the post-test, I wouldn’t have done nearly as well, without the teaching intervention” (p. 9). There is also the theory that if one repeats a test, one has a higher chance of doing better on it (Gall, Borg & Gall, 1996). Therefore, since all the students had to do the pre-test first, they may have fared better in the post-test, just because they had some practice with similar questions by taking the pre-test. Keeping this in mind, I asked Student 30:

> EV: ... So do you think that ... without the teaching intervention you could have done the post-test just as well as you did?
> S30: I am not knowing how I did on the pre-test as opposed to the post-test. I would say that I hoped I did better on the post-test because of the teaching intervention, and of course because of its ability to convey the same knowledge, but it reinforces in a different method, in different levels. (p. 45)

Student 30 makes a valid point that the students did not actually know their marks on the post-test, and therefore could not state with certainty that they had fared better in it. I did not let the students know the mark they got on the post-test, because I did not want to bias the students in any way, or give the impression that the post-test marks were important. I wanted the students to be as objective as possible when participating in the whole research sequence. In addition, I did not want any of the questions in the post-test circulating among the student participants. However, Student 30 hoped to have done better on the post-test due to the teaching intervention. This pedagogical intervention conveyed the substantive knowledge and procedural knowledge (i.e. Concepts of Evidence) using various human faculties, reinforcing both the students’ understanding and memory at different levels.

Theme two: Grappling with the APL. This was the first theme from the student interviews and explored what the students went through as they progressed through the research sequence shown in Chapter Three: that is as they read and grappled with the APL; answered the pre- and post-tests of the evidence survey; and participated in the interviews.

The first finding: Student experiences from reading APL. The experiences the students went through as they read the APL constitute the first finding. About two fifths of the students declared that they enjoyed reading the APL, while a lower number (about one third) agreed with reservations: some of these reservations are discussed below. For example, Student 6 who was not taking biology that year, had forgotten some of the words and terms necessary to understand the APL, and therefore needed extra help. Student 7 found the APL a little dry. Student 14 enjoyed reading the APL but “some of it was hard to understand and read ...” (p. 65). Student 19 did not receive the meanings of the words in the version of APL sent, and this was the only student to whom this happened. Student 21 commented that the APL was a bit too long, while Student 22 said the APL was okay, but the video was more enjoyable, because it was easier to follow when someone was talking. Student 38 remarked that the APL “wasn’t too entertaining, but from an academic point of view, I think I enjoyed it” (p. 15). From the interviews it was apparent that Student 38 valued very highly the academic experience of participating in my study, while Student 50, who enjoyed the APL at the beginning, found it a little boring by the end. Being forthright, Student 52 claimed, “It was kind of interesting, yes. I liked the meanings here. That helped a lot. Without it I would have found it really complicated” (p. 37). Student 57 described how the first reading was difficult, but it became easier as the readings were repeated several times:

> The first time round, it was interesting what they were getting at. I found that very interesting, but very difficult at the same time. It was very, it took a long time, of going back and trying to figure out what they were saying and stuff. that I read it, it was easier, it got easier and easier every time, but it was very interesting But it was the first time through so it was pretty difficult, but the second one and a few times after for sure. (p. 35)

It is noteworthy that this student found the APL “very interesting for sure” (p. 35).

Of the 55 students who participated, one in ten did not enjoy the APL. A balanced view about the APL was offered by Student 29, who said:

> It took some close critical reading, crash reading once or twice. It wasn’t too bad, because I hadn’t seen papers like that so far, in my research or in my classes. What I’ve read were a lot more in depth, so this was a kind of nice and [the best part was having] the definitions on the side. So it wasn’t too bad. (p. 34)

Student 29, who had prior experience reading journal articles, found the APL “nice,” especially since there was the added bonus of “the definitions on the side.” Well over a third of the students found reading the APL an enjoyable experience, though some had reservations such as the article being too long, difficult or dry, as discussed above.

The second finding: Intriguing parts of the APL. On examining what the students experienced as they went through the whole research activity, it was vitally important to know what they found most interesting in the APL. There were a number of contenders for the most exciting, interesting, and useful part. The most popular contestants were: the annotations or meanings provided; the beginning or introduction of the APL; and the interesting applications of this work. Some said they enjoyed the introduction because it was easy to understand, and others declared they “liked the beginning where they explained like why they were doing it” (Student 41, p. 48). The most useful parts of the APL, according to half of the 55 students who participated, were the annotations and meanings provided in the right column, while the other half voiced some other aspects of the APL as the most interesting. Some felt the APL provided a good background to the study, and some, such as Student 6, stated they had forgotten some of background as they were not currently taking biology courses, and enjoyed the APL “because it was such a pleasure and, the more and more I continued ...” (p. 7). A few students were very impressed with the “Materials and Methods” section because this really explained what they did and how, which they found interesting, revealing, and down-to-earth. On being asked whether the annotations and meanings in the APL were helpful, Student 31 replied “Those definitely helped: the explanations on the graphs were of great interest” (p. 51).

Student 43 found it particularly interesting to learn that molecular beacons could be used to detect something saying “I thought the fact that the molecular beacons, that you could actually use them to detect something” (p. 6) happening in the DNA. The same sentiment was echoed by Student 44, who said:

> They were talking about, way back here, what they used, this, it was spectrophotometers and I thought that was really interesting that they have actual instruments because I don’t think that they used a spectrometer that was really nice. Just like how chemistry and genetics came together, two of the hardest courses out there and the stuff was so interesting ... (p. 14)

Student 44 was fascinated by the use of spectrophotometers in this paper, and the interdisciplinary nature of this research which is perched on the frontiers of chemistry and genetics. This enthusiasm wore off on Student 46, who categorically stated, “Yeah, I don’t know if I enjoyed any part at all. It was okay, that’s it” (p. 31). A very similar stance was taken by Student 53 who stated this about the APL “... not really what I’m interested in” (p. 47). Student 49 expressed interest in the scientific methods used to detect damage in the base-pairs, saying:

> Just the methods that they were using to detect the damage, you know the damage to the base-pairs. It never would have occurred to me that you could do it like that, and so it just kind of opened up your mind as to just how these scientific methods work. (p. 7)

Appreciation for the practical scientific methods was also expressed by Student 51, who remarked, “I didn’t know how the molecular beacons have a light phase and dark phase: that was interesting” (p. 31). It was a pleasure to take note of what Student 54 said:

> S54: I enjoyed the graphs, the graphs showing like the comparisons between two different molecules like the ....
> EV: Okay, which of the molecules do you have in mind?
> S54: Just like the molecules such as the RNA and the DNA. (p. 7)

This student enjoyed the graphs comparing DNA and RNA. The fact that Student 54 could understand the differences between these two graphs and could connect this to the functioning of these two nucleotides (i.e., DNA & RNA) definitely demonstrated a superior level of understanding of the theory behind the graphs. This student found these comparisons most exciting and enjoyable.

When asked about the most exciting part of the APL, Student 57 replied:

> S57: Like, I would never have even thought of the fact that you could make something glow, DNA glow. Just the fact that they could figure out damage in base pairs through it glowing was just amazing.
> EV: It was like a light turning on. You couldn't miss it, could you?
> S57: It was fascinating, yeah. I had no idea that that was something that you could even do. (p. 35)

Student 57's expression of sheer wonder when describing the glowing of the molecular beacons as they combine with strands of DNA and are turned on, is refreshing, and suggested that, on the whole, some extremely enjoyable and poignantly memorable learning took place in this context.

The third finding: Genre of the APL. In tracing what the students went through during this research activity, it was important to know what they perceived as the genre of the APL, because their views on this matter revealed the third finding. "Genres are text types defined by function, socio-cultural practices and communicative purpose" (Baram-Tsabari et al., 2005, p. 404). Roughly half my sample of students claimed that they felt it was written in a descriptive or experimental genre, like an explanation. Approximately a third of the students thought the APL was written as a combination of two genres. Of these, half thought the APL was a combination of descriptive and narrative genres, while the other half opted for descriptive and argumentative genres. This is different from the accounts in the literature, where the APL initially developed by Baram-Tsabari et al. (2005) and Norris et al. (2009b) were claimed to have the argumentative genre. Their APL seemed to belong to just one genre while mine was perceived by students to be mixed, that is as a combination of two genres. Some students classified the genre of the APL by some different but original terminology calling it “Informative,” while others termed it “Scientific genre,” or “Scientifically descriptive.” Student 24 described the genre of the APL as both “objective” and “concise,” and claimed that the authors stated their opinions without bias. “I’d say narrative with explanations all the way,” (p. 51) was the view that Student 31 voiced about the genre of the APL. One of the main reasons for their importance is that diverse genres appear to evoke different types of mental processing (Alexander & Jetton, 2000). These varieties of mental processing are especially evident in the research on interest and importance specifically carried out by Schellings, van Hout-Wolters, and Vermunt (1996).

The fourth finding: The evidence survey: Positive student experiences. After having gone through the various stages of the evidence survey and completing them, the students must have noticed some positive aspects, and remembered some of the negative aspects as well. The positive and negative aspects would have contributed immensely towards their experience of participating in this research activity. In this finding, therefore, the positive aspects will be examined in more detail in order to respond to my first research question which entails examining the experiences of students as they read the APL, and went through the various stages of this research activity. Roughly a tenth of the participants said it was the novelty of the experience which they cherished, and found to be “pretty cool.” They had never done something like this before. One student remarked, “It was fun actually. It was a new experience, something really different, you know” (Student 27, p. 19). Some student participants really enjoyed the experience and described it as “good,” “interesting,” “positive,” and “I learnt a lot” (Student 8, p. 25). A tenth of the students voiced the opinion that the evidence survey was not too intense and that taking it added to the fun and excitement of taking part in this whole activity. Two students really liked the electronic access, which meant they could do the evidence survey at any time and from any place suitable for them; this tremendously enriched their encounter with this research activity. Some participants also had answers for the "Other" category and these will be examined in more detail below.

Some students, such as Student 6, found the evidence survey "hard," while others were quick to remonstrate, adding, "It wasn't too hard, it wasn't too difficult kind of thing, is what I think about this" (Student 10, p. 37). In short, there were students on both ends of the spectrum as far as considering the survey "hard." Some commented on the length of the survey:

> S11: It was long, but I think it is important.
> EV: Why do you think so?
> S11: You cannot just jump to conclusions on things. You need help and you need to be able to tell the average person in society, there is a reason why this and this and this is that. You have to understand and that's your part. (p. 46)

It is significant that Student 11 sees the need to communicate what was learnt to the "average person in society:" this is an attempt to make the public scientifically literate. This shows that this student perceives how learning about content knowledge and the Concepts of Evidence in the APL can be applied in the real world to promote scientific literacy. In order to achieve such goals, as Student 11 points out, you yourself have to understand, and according to this student, "that's your part." In other words, it is this understanding that Student 11 is trying to achieve presently, and hopes to pass on to the public. This is a commendable goal for someone who has just been introduced to the Concepts of Evidence.

Others took a more care-free approach and liked the way they did "alright" in it, without any pressure to perform and achieve. "I liked the way everything was spaced out, and especially because of the fact that you did alright, and it was fun doing it at the same time, and you feel like you are getting something out of it" (Student 12, p. 52). It is encouraging to note that students felt they were gaining something from this exercise. For instance, Student 13 remarked how the survey questions provided “food for thought” and the chance to do the thinking and come back to answer the questions. Student 13 liked this on-line interaction, and really appreciated the fact that one was expected to “think” and not regurgitate facts from memory. This student really liked the format of the evidence survey as well. Student 14, on the other hand, found the evidence survey interesting because it reinforced what one was already learning in biology and chemistry at that time.

Student 29 listed three points as positive aspects of the evidence survey: the video, which was really helpful in explaining and summarizing the paper; the diagrams embedded in the video, which were illustrative; and the pace of everything in the survey. Several students named the video as the most beneficial part of the evidence survey, asserting that the visual and the auditory stimuli in combination reinforced their ability to understand and retain what was being explained. One student claimed that through this activity one came to know a new technology, which one hopes to put to use some day. It is “so sweet you can do that with your DNA” (Student 42, p, 62).

Student 58 was appreciative of the helpful attitude of all who conducted the evidence survey:

> ... and everyone was very accommodating and willing to help and it was a great opportunity. It represented itself as a great opportunity for students and I found that very, ... I guess ... but kind of thoughtful, in a way, it was nice to have people thinking of the students. (p. 51)

Student 58 perceived it as a great opportunity being created for the students and appreciated such an effort to help students understand concepts which will help them engage better and deeper with science. Student 58 found this opportunity a thoughtful gesture promoting the welfare of students doing science in their first-year undergraduate studies. For Student 60, the biggest positive aspect was that “I get to have a peek out of someone’s data on a PhD!” (p. 68) and that, according to this student, was quite rewarding in itself. Student 47 added that the evidence survey helped interpret data better, and complimented it, saying, “It was more professional than what I have had to deal with so far.” The final positive aspect given by Student 47 was, “It gave me something which will always be with me” (p. 39). I hope the experience that Student 47 had while participating in the survey will enrich this student, and be cherished in the days to come.

The fifth finding: The evidence survey: Negative student experiences. The negative aspects of the evidence survey may be what these students remember and may have the most impact on them. That is why these negative aspects are examined in more detail here and why they contribute to the fifth finding. Approximately one fifth of the respondents said that movements of the camera could be heard as a rustling noise in the video, and that it was an annoying distraction for them. Student 45 said one negative aspect was that technology is great when it works, but frustrating when it does not function as expected. This student referred specifically to the incompatibility problems between Hotmail and Survey Monkey. This did cause some issues which led to minor delays but were, on the whole, quickly resolved. Approximately half the students could not really think of any negative aspects of the evidence survey, while the remaining students chose “Other” as their response, and they raised a variety of issues which are considered below.

Four students found it discouraging that though they read the APL, they failed to fully understand it. Student 1, for example, said that it took a while to realize that one had to note down a few points as one read the APL and refer back to them as one answered the questions. Some found the APL long and the whole survey time-consuming, taking longer than they had expected. This whole activity took away time which the students could have spent on their own academic studies. The negative part according to Student 20 was the questions, as she was unsure of what was expected of her. "I don't know what they're looking for here, and I don't know if I want to, but I get kind of confused with that" (p.37). Student 23 was of the opinion that the performance on the pre-test would have been better if the teaching intervention had been watched prior to attempting it. By the time this student got to watch the video, however, the pre-test had already been submitted. This is not a negotiable issue because the pre-test is intended to measure the initial understanding that the students have before any pedagogical intervention. It is vitally important, therefore, that the video is watched after the pre-test has been submitted, and this is exactly what is done in the evidence survey.

Student 25 was of the opinion that the five parts of the whole activity should be more closely spaced time-wise, and should be covered within a shorter period of time. Doing each part of the activity separately seemed a negative aspect at first, "but when everything added up it kind of makes sense, that's it" (p. 43). One student (Student 43) felt overwhelmed with the numerous diagrams, and was convinced that the emphasis should be on getting the overall picture. Finally, the timing in which the research occurred was not ideal, according to Student 60, who said:

> Maybe just the timing in which it took place in, I mean, it was kind of hard to really put the effort to meet all, just into understanding the article and figuring it all out while doing up the labs and stuff as well. (p. 68)

When I asked when would have been a better time, Student 60 mentioned the beginning of the year as a more viable possibility.

Theme three: "Wrestling with Concepts of Evidence": The views of students. This was the third theme and addressed my third research question concerning the views of first-year undergraduate students about introducing Concepts of Evidence as part of the curriculum in their first-year chemistry courses. The Concepts of Evidence which were not explicitly dealt with or even mentioned in the APL, are the building-blocks of the "how to do" of science as well as scientific investigations. The participants in this research activity had to struggle with these somewhat familiar concepts dealing with the usage of procedural knowledge in science and develop their own understanding of the importance and relevance of these concepts in the APL. This was a struggle and hence the name given to this theme “Wrestling with Concepts of Evidence” is very appropriate because it conveys how students grappled with these ideas and how the development of their understanding was in a state of flux during the course of this research activity. Sometimes, students brought with them some initial misconceptions that were corrected as the research activity progressed especially through the pedagogical intervention which entailed watching a video. After exploring the five findings, there will be a description of how they construed to this third theme.

The first finding: Selecting the most important Concept of Evidence. The opinions the students had about which is the most important Concept of Evidence constitute this finding. Of the 55 students who participated, half the students believed that Concept of Evidence #12, which is “Design: Validity, Fair Tests and Controls,” was the most important one for one basic reason discussed below. Roughly 25% of the student participants chose Concept of Evidence #19, “Detecting Patterns,” as the most relevant one, without which the meaning of experiments could not be deciphered. About 20% of the students regarded Concept of Evidence #5, “Instruments: Calibration and Error,” as the most significant concept, vital for making practical measurements, especially those using instruments, while a small percentage of the students regarded Concept of Evidence #11, “Design: Variable Structure,” as the most important.

The reasons the students offered for making their own choices of the most important Concept of Evidence varied widely. The most cited selection (selected by half of the student participants) was Concept of Evidence #12, and what was explored next was why this was so popular. Student 24 claimed that only a fair test can make an experiment in science valid, and setting up a valid experiment is the first step in proper empirical design. The warning pronounced by Student 40 was "If the whole fair test is wrong, there's nothing you can do about it. ... Then I mean, you're just done" (p. 40), "because that's what makes an experiment valid" (Student 7, p. 17). Another student underlined the importance of the fair test saying "Because if you don't have a fair test then even if there is a relationship, it's not going to be that one, because it's not going to actually happen" (Student 21, p. 7). According to Student 28, Concept of Evidence #12 is the most important one for the following reasons: first, then you will know that you are actually testing for the right thing; second, if you have a control, then you will know that other things are not interfering; and third, if uncontrolled, these other things will make you misinterpret your results (p. 27).

The second most popular choice (chosen by almost a quarter of the students who participated) was Concept of Evidence #19, "Detecting Patterns," and various reasons were afforded for this which will be looked at next. "I think detecting patterns and coming up with relationships is probably the most important," said Student 4, because "you know when you are conducting, like, a test, if something keeps happening, and if you get a pattern, you know that something is going on" (Student 6, p. 10). When patterns are detected "Then it really shows that we are making progress somewhere, and shows some sort of correlation, as in this case, between temperature and fluorescence intensity" (Student 11, p. 45). Finally one student had this to say about detecting patterns:

> S58: Because the more you see patterns the more and more you know you are heading in the right direction with your study.
> EV: How do you know? Why can't it be a wrong pattern? Maybe you're getting a wrong pattern.
> S58: Well that's true, but I mean, if the patterns are coinciding with your hypothesis, then something's got to be right. (Student 58, p. 52)

If the patterns you get coincide with the predictions of your hypothesis and the hypotheses of leaders in your field of study, there is a good chance that you are heading in the right direction, exploring the unknown. Data will make a lot more sense if you can detect a pattern in them.

The third most popular option (i.e. the choice of 15% of the student participants) was Concept of Evidence #5, “Instruments: Calibration and Error,” and the various reasons offered for this will be examined next. “Irrespective of hard work and being well-organized,” commented Student 27, “if there is calibration error or some other error, it’s not going to work. ... You’ll have some craziness” (p. 20). Another student insisted that “Zeroing is important for accurate/reproducible results” (Student 29, p. 36).

Only four students opted for Concept of Evidence #11, “Design: Variable Structure.” Next, we will examine what they had to say in defence of their least popular choice. According to Student 54, the most important factors to consider in the design of an experiment are the three different types of variables – the dependent variable, the independent variable and the controlled variable. Without these, one would not be able to establish what these molecular beacons are affected by, and what does not affect them. Student 17 was quick to point out “It’s important to have the controls that you are going to have, and have the variables that you are going to have, and keep that the same throughout the whole experiment” (p. 12). Student 13 commented further on this requirement, saying:

> Well ... I think ... there’s the two variables there. I think you say manipulated and the dependent variables. ... you will understand, you’ll always have to have the manipulated variables and controls, and the dependent variables of course, to see what’s up. And then, but of course, the correlation is really important, because if you follow that track. But if you’re skewing results and cannot measure up, then you are not getting what is in reality. (p. 60)

When carrying out experiments or quasi-experiments, it is important to change, observe, and control certain variables. Once this is done, the relationship between the variables must be studied, and if any correlation exists, it must be detected, established, and reported as a result. The absence of correlation is in itself a result that can be reported.

The second finding: "Fair test" defined. The ways in which the students responded initially to defining a fair test are categorized into three levels. A quarter of the student respondents gave a definition that was low in accuracy, while half of them gave a definition that was of medium accuracy; and 10% were fully accurate in their definition. The low accuracy definitions often did not mention variables, and did not recognize that variables could be classified into different types in a fair test. These answers often just claimed that the conditions had to be maintained the same, but did not explain how exactly that was achieved in an experimental setting. Some definitions which were low in accuracy also claimed that a fair test implied the experiment had to be repeated more than once. Other answers in this category did not identify the various types of variables accurately. The main theme that emerged in these low accuracy definitions was that in a fair test the experiments had to be done the same way and some conditions had to be kept constant. The answers in the medium accuracy category showed an awareness of the existence of different types of variables. They usually identified the dependent and independent variables accurately, but failed to mention the fact that in a fair test one independent variable had to be varied at one time. A fully accurate definition of a fair test included the fact that in the design of a valid experiment in science, the dependant variable (which may be one or more) is observed, and only ONE independent variable is varied at a time.

The third finding: Classification of variables. An awareness of what the variables are in an experiment, and the ability to classify them, are both vital for understanding the role of Concepts of Evidence in the design and performance of an experiment. The students were told that absorbance and fluorescence intensity are being measured in Figure 6. They were then asked what sort of variables they are, and were required to classify them. Half of the student respondents chose the right answer, classifying the variables as either dependent variables (which was the more popular term), or as responding variables, which is a synonym. This was encouraging, showing that over 50% of the students were aware that variables existed and knew how to classify them in an experiment. These are basic understandings of procedural knowledge (or Concepts of Evidence) required for the process of designing any valid empirical activity. The dependent or responding variable is the one that is observed and watched in an experiment. Only one in ten students classified dependent variables incorrectly as independent or manipulated variables. One student called them “experimental” variables.

Students were pointed to the “Time of Irradiation” as it appeared on Figure 6, and were required to classify that variable. Of the 55 students interviewed, 38 correctly placed “time of irradiation” as an independent variable - which was by far the more popular term - or as a manipulated variable - which was the less common usage. This is the variable that is changed or manipulated in an experiment. There were only three students who wrongly classified it as a dependent or responding variable. I think the students fared better in classifying this variable, partly because of the experience they had in the previous example with dependent variables.

The fourth finding: Manipulating independent variables. Once the students had recognized dependent and independent variables in an experiment, in order to probe the understanding of Concepts of Evidence even further, they were asked how many independent variables are manipulated at one time. This was a well-answered question, in which 64% of the respondents said only one independent variable was changed at one time. Only two students chose the “Other” option; Student 13 could not answer the question asked, but gave evasive, unrelated, and unclear answers, until finally selecting “one” as the correct answer. Student 23 gave an uncertain, “usually one,” as the answer: it is not just “usually,” but it is “always” only one independent variable that is altered at one time.

There were two students who described in a very lucid fashion when an experiment in science is valid. Student 14 stated how, in order to make sure that an experiment in science is valid, one has to:

> Make sure that it’s valid and that you’re given anything, you can’t just prove it actually true or false. You have to ... like doing all those controls. Make sure that what they found out about in this experiment is actually true ... (p. 67)

This student failed to recognize that it is when a fair test is established in the experiment that is when only one independent variable is altered at one time, that the experiment becomes valid. Student 16 capped this line of thinking, stating that an experiment in science is made valid by “making sure your findings were the actual true findings” (p. 6). How does one ensure this? The best way to ensure the validity of an experiment is to design it as a fair test, where only one independent variable is changed at one time, and all other independent variables are held constant.

The fifth finding: Varying two independent variables simultaneously. Once it was established that in a fair test only one independent variable can be changed at one time, the students were next asked whether they could alter two independent variables - the time of irradiation as well as the magnesium concentration – in the same experiment. There was a resounding “No” from 38% of the students. One student (Student 37) wrongly answered “Yes,” saying “You probably can do the procedure, but it will be more difficult” (p. 9). The fact that this error undermines the very validity of this experiment, rendering it invalid, does not occur to Student 37.

Three students chose “Other” as the category of their responses. Under this category, Student 19 affirmed that one could not change the time of irradiation and the magnesium concentration in the same experiment, “because you change one at a time” (p. 30). No further explanation was given by this student as to why it is vital to change only one independent variable at a time. After initially asking Student 51, I explained that if an experiment is to be valid or have validity, it must be a fair test. I proceeded to define a fair test as an experimental design in which only one independent variable is varied at one time. I then asked whether, if I wanted to find out how the magnesium chloride concentration affects the fluorescence intensity, I could use the same experiment, and have two independent variables. Student 51 responded with an emphatic “No.” I continued:

> EV: ... So you can have only one independent variable at a time. If you wanted to find the effect of magnesium chloride concentration what would you have to do? Do ....
> S51: Another experiment
> EV: Another experiment, right. So you do a second experiment, keeping the time of irradiation constant and varying the magnesium chloride concentration. (p. 33)

Student 51 has obviously taken the answer a step forward, and reflects a deeper, more practical understanding of a fair test. I introduced Student 57 to the scenario saying:

> EV: ... suppose here I’m interested in the effect of time of irradiation on fluorescence intensity and absorbance. I may also be interested in the effect of magnesium concentration on fluorescence intensity and absorbance. Do I do that in the same experiment?
> S57: No. You do two separate ones. (p. 38)

Student 57 was quick to remonstrate that two separate experiments are necessary if one wants to examine the effect of the time of irradiation and study the impact of the magnesium concentration on the fluorescence intensity and absorbance.

All the findings explored in this third theme were about Concepts of Evidence, especially the fair test, and the background information required to understand them. A lot of effort was put into outlining and explaining all the intricacies of a fair test so that students would understand the concept. That is why all these explorations contributed to the third theme, “Wrestling with Concepts of Evidence.” There is an aspect of “struggling to understand” that is conveyed by the word “wrestling,” and I think this conveys accurately what many of the students experienced as well. The students had to read the APL once, and then come back and read a second or even a third time, and then struggle to really make sense of what was happening in some of the experiments described.

Theme four: Challenges for attaining substantive knowledge. This was the fourth theme that arose from the student interviews. Five findings contributed to this theme, and they are discussed individually below. I begin with a description of how I arrived at this fourth theme.

The five interview questions in this theme and the resulting sub-sections were based on substantive knowledge in the Adapted Primary Literature (APL). The questions raised were on a wide variety of content topics ranging from what upward and downward arrows depict in a figure; anomalous results and what can be done with them; what sampling is, and recognizing its presence or absence in an experiment in the APL; a control and how to set one up in a particular experiment portrayed in the APL; and the scientific purpose of the whole series of experiments traced in the APL. Since the five questions were based on basic content knowledge essential for recognizing and understanding the Concepts of Evidence nested within the APL, all the findings dealt with how students responded to these questions, and shed more light on the challenges they faced in understanding the substantive knowledge.

The first finding: Dealing with outliers or anomalous data. The first finding was the name one gives to data which deviate from the general trend. Almost 50% of the students called them “outliers,” a term especially familiar to students who had taken a course in statistics.

Although this usage was the more mathematically correct term, in science they are also known as anomalous data, and only two participants out of a total of 55 labelled them as "Anomalous Results." Only one student managed to get both terms, "Outliers & Anomalous Results," which was the most inclusive answer. During the interviews, I did not specifically ask the first 11 students for the term, referring to the data that do not obey the general trend of data in an experiment. However, later I felt this should be included. This is a good example of how the line of questioning developed during the course of the interviews. Nine out of the 55 students who participated chose the "Other" option as their answer. Six of those who chose "Other" said that data that deviate from the general trend are called "Errors" or "Erroneous Data." Three other respondents called them "Abnormal," "Extra Points," and "Interpolated Points" respectively.

The second part of this finding involved describing the two procedures that can be adopted when anomalous results are obtained during data collection. More than 50% of the students who participated said that the experiment should first be repeated, and if the anomalous results persisted on repetition, then they should be explained. A quarter of the students said that the results should be explained, while 18 students out of the 55 (i.e. 33%) suggested that the experiment should be repeated. There were a number of offers (5 out of 55) under the "Other" option. Each response in this category will be dealt with in detail below.

Student 8 said initially "Yes, I would recognize them as weird. Anyway there is an error. ... I guess I might do that experiment again," but did not mention anything about explaining the anomalous data. Student 12 explained that if there were multiple occurrences of anomalous data, the solution would probably be to redo the experiment, but added, "Aside from that I could just call it error, if I had just one outlier." When asked what could be done with anomalous data, Student 58's reply was "I would either, I guess, extend the points I guess would be one thing."

Student 58 was then asked to further explain the meaning of "extend the points," and went on to claim, "Well you go to, a logarithmic function, so that you are in a straight curve." When asked whether this statement meant that "those points will be brought closer to the curve," Student 58 replied, "I don't know, maybe," indicating a degree of uncertainty which this student could not clarify in spite of being requested to do so.

The second finding: Integrity in empirical science. Once the initial task of establishing that data which do not concur with the general trend should be called "outliers" or "anomalous data" was accomplished, the interviewer asked "What are the two things a researcher could do with anomalous data?" Half the student participants offered that the participants could either "repeat" the experiment or "explain" the anomalous findings. About half that number offered only "explain" and a higher percentage "repeat the experiment" as the appropriate answer. Interesting issues underscoring the importance of integrity when carrying out experiments in science arose when the interviewer asked "Would you sweep these anomalous results 'under the carpet?'": these issues comprised the second finding.

Student 18 responded to the above question with:

> If it is just one or something, it may not affect the results, and you might have enough data to not use it, but if there is multiple, then it might be if you ... that you waste the time on repeating the experiment. (p. 14)

According to Student 18, just one stray case of an anomalous result is not going to affect your results, but a multiple occurrence of anomalies needs to be noted, and the experiment repeated. It must be borne in mind that Student 18 considers this repetition of the experiment a waste of time.

Student 30's response to this question was that one could not ignore anomalous results. Therefore, one would have to repeat the experiment, by preparing fresh solutions. I then asked this student what could be done if the same results were obtained again. Student 30's response was that if that were the case, then one would not consider the results anomalous any longer, but one would need to examine whether there might be an error in the procedure, or whether the procedure in this experiment is interacting in unexpected ways that have not been accounted for. It is noteworthy that Student 30 regards overlooking anomalous data as equivalent to taking the tremendous and hazardous risk of "ignoring potentially ground-changing results." One could, thereby miss out on lifetime opportunities for making ground-breaking discoveries.

Therefore, one needs to take an ethical stance in science. The gist of what the students had to say on this topic of experimental integrity was that if one gets anomalous data, one has to find some way of justifying them, or explaining them away. It would be easier to "sweep them under the carpet," but in science one has to be accountable. For example, you may do the same experiment tomorrow and find that things have changed and there may be a reason for it. It may be you are onto something really important. Sometimes, advances in science happen because of such anomalous results, and one has to watch out for that. That is why when you get anomalous results, it is important that you explain them or re-do the experiment to ensure that the results are real and reproducible, and not an artifact.

The third finding: Identifying "sampling" in experiments. The very first question the participants were asked, as an introduction to "sampling," was whether they knew the meaning of this word. More than half of the students interviewed did not initially know the meaning of this term. I explained the meaning to them, in great detail, using the example of the experiment employing the irradiator, which was described during the teaching intervention. In response to the question, "Is there sampling in Figure 6 in the APL?" 87% of the students answered in the affirmative. Those who chose "Other" as their option, and their responses, are considered individually in the discussion below.

Students were next asked why they either thought there was sampling, or were of the opinion there was no sampling. The reasons the students provided for the affirmative and negative answers were combined to constitute the third finding. Student 23 answered the question “Why?” by saying “I think there would be sampling here, because the dots seem to be all around certain lines along the x-axis. Like there are a lot around time ‘0’, and a lot around time 50, not really much in between.” There was no further explanation as to why the close proximity of the dots indicates that there is sampling.

Student 29 thought there was sampling, and categorically supported this idea by stating:

> Because they have to ... like you’re testing how irradiation affects the damage, the amount of damage. So you have to take a whole bunch of little samples so that you can get this trend line. If you just took that whole sample ... (p. 31)

In Figure 6, the reason that Yarasi et al. (2005) were able to obtain several readings at very short intervals, in times of irradiation, was that they took several aliquots from the same hybridization mixture.

When asked if there was sampling, Student 44 emphatically stated:

> You can’t expect to put a whole, like, all of your army into your experiment and then take it out and then have to reset it up and leave it in for longer time or you could just take a little bit out and then test it and then let the experiment continue on. (p. 11)

This student compared the experimental set up to an army. Student 44 recognized the importance of irradiating the hybridization mixture continuously, and that is a part of the reason that sampling is employed in this experiment. This student added that taking the whole hybridization mixture out for each test did not make sense, and reiterated “you know just time-wise, material-wise it doesn’t make sense.” Sampling, according to Student 44 is the most sensible option.

The students’ understanding was further probed by asking them to justify why they said initially that there was sampling or not. Most of the answers reflected some understanding of the fact that the whole reaction mixture put into the irradiator was the original hybridization mixture, and only small aliquots were withdrawn from this mixture, at various time intervals, to determine the fluorescence intensity and the absorbance: this meant there was sampling. There were a number of variations of this answer.

For example, Student 7 when asked why one would think there was sampling, maintained that the readings were taken at different times and this implied sampling. Student 8 justified the existence of sampling by stating "I think so, and therefore, I guess they're measuring samples here (pointing to the dark and the open circles in Figure 6), and those are the dots." Student 10 argued that there was sampling because "it is a new sample he takes every time." In the same trend, Student 16 explained, "Because in the entire figure there are so many points." Pointing to the dots on Figure 6, the justification for sampling provided by Student 18 was, "I think they'd have to be aliquots in order for you to get reliable results out of it." Student 19 defended sampling saying "... the individual dots which are individual samples being taken, but they're being sampled ...." The same argument was expressed by Student 20, who said, "Well you see the different points. The points each of those is a different sample that you are testing." Student 46 reasoned that sampling existed in Figure 6 because "Each of these (pointing to the shaded dots) is from the same vial." This is a very appropriate and succinct answer.

Student 51 initially said, "I'm going to say 'No' because there are different concentrations of the absorbance and fluorescence intensity." The discussion then continued as:

> EV: So my question is, Is there sampling in this experiment? What do you say?
> S51: My first answer was "No," but now I don't know...
> EV: Why do you say "No"?
> S51: Well, because I thought the concentrations were different. (p. 26)

The reason offered above was not true, because sampling can occur irrespective of whether there is a difference in the concentration. This is a misconception and, therefore, I pursued the matter further with more questions.

EV: Which concentrations?
S51: The concentration of the rU_{17}
EV: And ...
S51: Of the molecular beacons ... of the hybridization mixture
EV: They are different where? Between what and what?
S51: The buffer and the hybridization mixture. (p. 26)

When I explained that the concentration has no bearing on whether there is sampling or not, Student 51, changed the answer to "Yes," there is sampling, and defended it by stating, "Because you have as the time goes on ... it is not ... it is completely different from when you have to leave it in, then take it out, and then hold it in for a while longer." The impracticality of taking out the whole hybridization mixture, and then later putting it back into the irradiator, and re-heating to obtain the total time required, is what is alluded to above. I then asked whether, looking at the figure, Student 51 could provide a reason for stating there was sampling. The final defense Student 51 gave for sampling was: "Because the dots follow the curve and then make a line of best fit." This is a start, but does not fully explain the evidence for sampling. However, there seemed to have been a radical change in the students' opinion about whether there is sampling in Figure 6. The whole interview here is an example of how there was a misconception in the student's thinking, how it was corrected and how that then changed the student's view on sampling. This is an attempt, on the part of the student, to attain a conceptual change.

The excerpt below shows how Student 55 responded to the question, "Is there sampling in Figure 6?":

S55: Yes, doesn't there have to be?
EV: Or could it be each of these points is another experiment with a different ... the whole sample is being changed or is the sample being taken out of the original ......
S55: Sample being taken out of the original.
EV: Why do you say that?

> S55: Because it would be very time-consuming to do each one differently or you can do it all at once and you can measure the exact time in-between, say, you have to wait 50 minutes, and then you have to wait another 50 minutes. (p. 15)

The first response of Student 55 is loaded with meaning, and shows that this student has realized there is no better alternative in this case to sampling. Therefore, it is essential that sampling is adopted here. This student provides three good points to the question “Why?” First, it would be too time-consuming otherwise; second, the whole experiment can be done at once; and third, you can measure the exact time in between the various “scooping out” of the aliquots. These three points are noteworthy, and show that Student 55 has clearly understood not only the concept of sampling but its advantages.

The fourth finding: Setting up control experiments. Participating students had to decide what to call experiments that are done on the side, to shed light on the main experiment, and how to set up one such “control” experiment based on Figure 7 in the APL. These surmises, along with some basic substantive knowledge discussed below, comprised the fourth finding.

The students were asked whether they knew what “d” and “r” stood for. Out of the 55 students who participated, 30 recognized “d” as DNA while only 11 identified “r” as RNA. This shows that students are more familiar with DNA, perhaps because in their secondary school science and biology classes, they have come across this molecule and the role it plays in heredity: they may be less acquainted with RNA.

When asked specifically about experiments done on the side, 40% of the students said such experiments were called “controls”; 26% were not sure of the answer; while six students chose the “Other” option under the answers provided. Of these, Student 4 did not get the word “control” but defined it correctly as “No, you measure something that has everything but the sample in it.” Student 12 had a longer explanation of what a control is: “It could be just eliminating the possibility of some of the factors affecting the results. They are kind of out there. We know what they are. Anything else other than molecular beacons is to be controlled." This student realizes that variables that could skew the results have to be controlled, but does not fully define what a control is. Student 14 said, "I don't know what they are called, but they seem to be there ... just to reinforce the data that you have." The students gave the term "control experiment" a variety of names such as "confirming experiments," "secondary experiments," "supporting experiments," and "background experiments."

More than 10% of the participants explained the method of setting up a control experiment. The goal of the control experiment was to demonstrate that the fluorescence intensity observed in Figure 7 of the APL was actually due to the molecular beacons and not anything else. The best way to set this up was to have a control with DNA and magnesium chloride, which would have no fluorescence. Molecular beacons are then added to this control which would immediately start to fluoresce. The first line of reasoning suggested by about 50% of these respondents was to take the magnesium chloride out of the hybridization mixture. The fact that removing the magnesium chloride would not demonstrate that the fluorescence intensity was due to the molecular beacons had to be addressed and rectified by the researcher. This was done by showing that removing the molecular beacons instead from the hybridization mixture would yield the appropriate control. The most frequent answer was setting up a control with DNA and magnesium chloride. A few thought that the control could be set up with just molecular beacons. Only one student thought of setting up a control with just DNA.

Some interviews revealed that the students harboured certain misconceptions or alternate conceptions initially about how control experiments are set up. For example, Student 39 claimed that the control was set up by taking off the independent variable, which was the time of irradiation, in Figure 7. Controls are not set up by removing the independent variable: this is a misconception. One of the components of the original hybridization mixture is removed to set up the control, and in the case of Figure 7, it would be the molecular beacons. The best way to set up a suitable control therefore is to consider how to set up a hybridization mixture which fluoresces, and from that remove just the molecular beacons, which are the molecules of interest. If this is done, one will end up with a mixture containing all components of the hybridization mixture except the molecules of interest: that is the required control.

The fifth finding: Purpose of experiments in the APL. The fifth finding comprised the scientific purposes offered by the students for this whole series of experiments in the APL. In the interviews, 42% of the participants identified the scientific purpose of this whole series of experiments as finding the optimal conditions for and comparing the efficacy of the molecular beacon probes in measuring the photodamage in thymine and uracil oligonucleotides when exposed to the uv-light source. Nearly a half of the respondents chose the “Other” option under the answers provided: Some of these answers will be considered below.

Student 4 responded saying, “They were looking at ways to observe damage in DNA. ... detect the damage using fluorescence” (p. 10). This answer did not include the term molecular beacons. When I provided that, Student 4 chuckled, as if the omission had been a lapse of memory. When Student 5 was asked to identify the objective that Yarasi et al. (2005) had for working on this series of experiments, the reply was, “The effect of uv-light on molecular beacons” (p. 15). This was the closest Student 5 got to the right answer. When further questioned: “And what were they trying to do with these molecular beacons? ... Was there any hypothesis they were trying to test?”, Student 5 simply replied, “I am sorry, I don’t know” (p. 15). The scientific purpose offered by Student 11 was “Pretty much to see if photo light gives a direct reason why things are the way they are. Trying out experiments by having a control, trying wavelengths and all out" (p. 44).

Student 13 provided a very general answer focussed on applications, and described the objective Yarasi et al. (2005) were trying to attain this way:

> I know they were trying to see the damage the sun could do on the skin, and all that. So, I know they did different" variations" because there are different variabilities in the course of that, and whether what's going on is or is not so scientific on a molecular basis, just like how the different skin types, the different intensities and the different altitudes and that will affect uv-radiation, and the natural or tanning bed, and that kind of shows different variations of that. (p. 56)

This is a very broad answer which needs to be more precise. The purpose offered by Student 14 was, "They are trying to find the energy level which affects the DNA." When I asked what the authors were using for this, the response was, "Using the molecular beacons to test the fluorescence ..." (p. 65).

The initial response of Student 20 about the goal of this series of experiments was "they are looking at, like, photodamage, ... in thymine (T) and uracil (U)" (p. 35). From Student 20's answer, I drew out the fact that thymine is in DNA and uracil in RNA, but there was no mention of molecular beacons, and that is what I was trying to obtain when I asked for the "tool" that was employed. The response was, "They're using, like, the fluorescence intensity and absorbance. Is that what you mean?" (p. 35). I then reminded Student 20, those were the "methods" and not the "tool." As this student still seemed unsure, I pointed to the diagram of the molecular beacons in the paper (Table 1) and asked what those tools were. With that hint, Student 20 figured out the answer "... those were the probes, the molecular beacons." Sometimes it was important to ask probing questions, as shown above, and guide the thinking of the students so that they arrive at the best answers.

When asked to identify the purpose of this series of experiments, Student 29 replied, "They're trying to find damage in DNA or RNA but by very specific, by just one nucleotide sequence" (p. 34). Student 43 got the scientific purpose from the title of the paper, which I gave as a hint. I tested this student's understanding of each word in the title. Student 43 knew all the meanings but did not mention that these molecular beacons were the probe used to detect photodamage in the thymine and uracil oligonucleotides. This seems to be a common, recurring shortcoming in the understanding of the overall objective of the series of experiments discussed in the APL.

Analysis of the Interviews of Instructors

The nine chemistry instructors interviewed were from eight different postsecondary institutions within the province of Alberta. I was given access to the first-year general chemistry classes that these instructors taught. Urban as well as rural postsecondary institutions were included in the sample. The sizes of the postsecondary institutions chosen varied from small to medium to large. The interviews with the instructors were treated just like those of the students: they were coded and the information obtained was utilized to fill out the Analysis Forms. These were further studied to provide the main themes discussed below.

Theme One: The Instructors' views on the teaching intervention. I sent the teaching intervention, as an attachment, directly to the students when they had completed the pre-test on Survey Monkey. Since it was not mandatory, three-quarters of the instructors did not watch the on-line video which was the teaching intervention. However, when they were interviewed I asked the instructors for their opinions on the teaching intervention, and their views contributed towards the only finding discussed below.

The effectiveness of the teaching intervention: Instructors' opinions. The instructors' opinions about the effectiveness of the teaching intervention varied. Most instructors were convinced of the efficacy of the teaching intervention in promoting a good understanding of the Concepts of Evidence. The visuals and anecdotal explanations were lauded as being pedagogically sound in successfully linking empirical evidence to the underlying scientific theories. One instructor vouched that the teaching intervention would certainly further the students' learning of the substantive knowledge and the procedural knowledge (i.e., Concepts of Evidence) contained in the APL. The responses of roughly half the instructors fell under the "Other" category and these are explained on an individual basis below.

Instructor 4 was of the opinion that the students do understand the Concepts of Evidence, but had reservations about whether this was an "explicit understanding," and whether they were able to draw the necessary connections to their own experimental work. However, this instructor thought the students had grasped "the way experimentation is done" and the importance of "tying experimentation with the theories."

When asked whether the teaching intervention in the form of the video can make a difference, Instructor 5 responded emphatically:

> I think so. I like being in front of .... I like teaching live so, I think that's one way, but I don't think the only way, ... so having something different and especially this day and age of students, they are more video, audio .... so the idea of them, you know, watching something and paying attention to it is probably, they might find that more exciting than a live teaching. So, I would imagine it certainly was effective. Yeah. (p. 7)

Instructor 5 had no reservations about the effectiveness of the teaching intervention, because of the certain awareness this instructor had that these particular students would welcome a change and happily watch and learn from the video. This instructor surmised that these students would perhaps even prefer that to the live teaching that they are given in class. Instructor 6 was also of the firm opinion that the "teaching intervention definitely had a role in improving the understanding that students have of these Concepts of Evidence," and was quick to add that all teaching interventions improve students' understanding. Instructor 7 also commented very favourably about the teaching intervention, saying “I think it will make a huge difference. I predict that it will make a huge difference.” Instructor 7 added that despite the fact that there was only one, the teaching intervention will have tremendous efficacy in furthering the students’ understanding of both substantive knowledge of the APL and procedural knowledge (i.e., the Concepts of Evidence embedded in the APL). These positive comments about the efficacy of the teaching intervention were very encouraging. The first finding was based solely on the opinion the instructors had of the effectiveness of the teaching intervention. These varied opinions of the instructors contributed directly to this second theme, and swayed greatly in favour of the efficacy of the teaching intervention employed in this study.

Theme Two: Challenges in the content knowledge of the APL. This first theme from the instructors’ interviews examined their perception of the substantive knowledge within the APL and explored how challenging they found it to be. The content knowledge included in the APL is interdisciplinary in nature. This posed challenges for those who were not familiar with how knowledge overlaps in the fields of biology, genetics, photobiology and photochemistry. In this study the understanding that instructors hold of Concepts of Evidence was tested through the medium of the APL. Therefore, it was important to grasp the substantive knowledge in the APL in order to fully appreciate the role that the Concepts of Evidence played in controlling the details of how these experiments were carried out. When interviewed, the instructors raised two main concerns about the content knowledge of the APL: First there were unjustifiable omissions and second some of the details were elusive. These two concerns comprise the two findings under this theme.

The first finding: Unjustified omissions. The first omission the instructors were concerned about was why arbitrary units were assigned to the fluorescence intensity (FI) in

Figures 1, 2, 3, 5, 6 and 7 of the APL. Some instructors found it disturbing that the authors of the primary literature (Yarasi et al., 2005) assigned arbitrary units for the fluorescence intensity without any explanation as to why. These instructors experienced challenges as they attempted to justify this allocation of arbitrary units. A quarter of the instructors said that the authors of the APL were only interested in the change in the fluorescence intensity (ΔF). Another quarter said getting an absolute value was not important, since this was only a relative scale, and what was required was only a relative fluorescence intensity. Some of the respondents were unsure of the answer, while a few gave "Other" as their choice. The observation made by Instructor 6 was that "... this is a trend they're looking at, and it is more of a general thing." The most laudable answer, however, was provided by Instructor 7, who said, "The arbitrary units are probably just a function of the fact that they are 'ratioing' their signal to the baseline, and so it's comparing their signal versus a fender baseline." If the fluorescence intensity has arbitrary units assigned to it, it is probably a ratio, and units do not matter. When I investigated this further I learnt that fluorescence intensity has the property of varying with time, and of being difficult to reproduce exactly, a drawback accepted by scientists who work in this field. Yarasi et al. (2005) had conformed to the standards practised in this area of photochemistry and photobiology.

Most of the instructors were disturbed by the second omission that it was unclear whether the points plotted on the figures in the APL were from single or multiple trials of the experiments. I believe this should have been clearly specified in the primary literature by Yarasi et al. (2005). What I learnt here was the importance of specifying all the conditions used, putting in error bars, and articulating the number of trials done when carrying out experiments. Failing to take these measures is an unjustifiable omission because it casts doubts on the validity and reliability of the investigations reported in Yarasi et al. (2005).

Among the interview questions posed to the instructors was one examining the criteria used in choosing the targets which bind to the molecular beacons. The targets chosen were the simple oligonucleotides - dT_{17} and rU_{17} – which would “behave” like DNA in their experiments. Why were these two oligonucleotides selected? There is no answer to this question in the primary literature by Yarasi et al. (2005): a third omission that the instructors found troubling. Two instructors thought that Yarasi et al. (2005) knew that thymine and uracil were the most photochemically active bases in DNA and RNA respectively. Instructor 2 thought that these authors (Yarasi et al.) wanted a very well-defined, simple sequence where they could be sure damage would occur, and where they knew the damage is predictive. For Instructor 7, the criterion used was simply that this was the most controlled and predictive experiment. According to Instructor 1, the reason for selecting dT_{17} and rU_{17} was that these do not occur in normal DNA/RNA sequences. Instructor 2 offered two answers: first, these two oligonucleotides were well-studied and perhaps were cheaper; second, “they also may be the ones that are associated with the kind of damage that takes place.” Instructor 6 offered a practical reason: that these researchers had to start somewhere. This is their preliminary work, and they may go to other oligonucleotides in later experiments.

The second finding: Elusive details. One of the interview questions to the instructors, was based on Figure 2 of the APL, and examined what the similarity in shape between the heating and cooling curves tells one about the underlying structure of the molecular beacons. Since this involved linking simple physical observations to a complex underlying structure one had to be very specific with the content knowledge employed. Half the instructors were of the opinion that the similarity in shape was indicative of the fact that the changes occurring during the heating are reversible during the cooling phase. The other half chose the “Other” option, which had four interesting responses. Instructor 4 was of the opinion that this is a folding curve; "... it is folding: It folds back to the same shape." Instructor 5 said, "I think it's that it behaves the same whether it's warmed up or cooled down, so that it behaves the same at all temperatures." The latter part of this statement was deemed untrue by Instructor 6, who stated, "... the thing is that it has to follow the same pattern." Instructor 8 maintained that, "as you said, they are superimposable, like it didn't do anything." If two curves are superimposable, it need not be because nothing is happening, but because the events occurring are reversible. In other words the changes that occur when heating the molecular beacons are in the exact reverse order to the changes that take place when they are cooled. Consequently, the curves in the forward and the reverse order are superimposable. The instructors experienced a lot of frustration in relating the physical observations made of the molecular beacons to the changes that are occurring in these probes at the molecular level. Envisioning such complex relationships demands a thorough understanding of the content knowledge, without which such details are elusive.

From the interviews it was evident that the instructors were able to answer much harder questions than the students. This may be because the instructors are more highly qualified and have had a lot of previous experience in reading primary literature. However, similar to theme one from the student interviews which described the challenges they faced upon trying to attain a good understanding of the substantive knowledge in the APL the instructors also had hurdles to overcome in order to fully grasp its content knowledge. For example, it was challenging to link the underlying structure of the molecular beacons to how they behaved when they were heated and then cooled (i.e., their heating and cooling curves). I learnt this was a good example of how in western science physical observations are always related to the changes that are occurring at the molecular level.

Theme three: "The thinking gets kind of lost in the content." The title given to this theme is a significant quotation from Instructor 5. It suggests that the thinking associated with the Concepts of Evidence (or procedural knowledge) gets somewhat submerged in the substantive (or content) knowledge. One of the most interesting issues that arose from the interviews of instructors was: "Are Concepts of Evidence to be taught explicitly or are students to be left to struggle with them?" This theme deals with "How Concepts of Evidence influence the teaching of chemistry and the role they play in the teaching of this fascinating science discipline." There are three findings under this theme and all three were derived from questions posed to instructors geared towards advancing the constructivist pedagogy.

The first finding: Teaching chemistry: The role of Concepts of Evidence. The first finding is focussed on what the instructors perceived as the role of the Concepts of Evidence in the teaching of chemistry. Nearly a tenth of the instructors felt that their role was to stipulate a consideration of the factors controlling the collection, measurement and handling of data. The answers provided by the instructors are discussed individually below.

Instructor 1 provided a brief answer to this question. The sole reason provided for introducing these Concepts of Evidence was "... this is a good thing because you are going to ask students to think about" (p. 3). This instructor went on to add that in science you may think you know something, but when someone asks "Why? How? Prove it," you cannot (p. 3). You are stuck. "You've got a bit of egg on your face on that. ... But that's the way science works" (p. 3).

Some of the more experienced instructors were very enthusiastic about conveying the paper on Concepts of Evidence to the professors in charge of teaching the sampling course in chemistry at their institutions. They expressed regret at not having been taught anything at all about procedural knowledge when they learnt chemistry under the older curricula. Obviously, addressing Concepts of Evidence in chemistry classes is a recent phenomenon. I was impressed that one instructor insisted that Concepts of Evidence are being addressed “piecemeal” in the present undergraduate chemistry curriculum as well, though they are not termed as such.

When asked about the role of the Concepts of Evidence, Instructor 2 replied:

> I think it’s quite valuable. In the olden days, when I did my studies, we never did. ... We did discuss a little bit of this in our sampling course, but I am going to refer that article to the present teachers handling that course, because I think it’s a very good article that will relate for undergraduate students. I think that will give us some ideas. (p. 7)

Instructor 2 was definitely of the opinion that there was value in teaching these Concepts of Evidence. It is noteworthy and very encouraging that this instructor’s first reaction was to think of referring the paper by Gott et al. (2003), on Concepts of Evidence, to the teacher teaching the sampling course in that institution. Instructor 2 felt that this article was a good one and that it would relate to the undergraduate students. This more experienced instructor was quick to acknowledge that procedural knowledge was never addressed in chemistry classes in their days – “not one iota” (Instructor 2, p. 7).

The third instructor claimed that these Concepts of Evidence can be delved into more deeply. Reproducibility, for example, can be taught very superficially, but can also be treated at a much deeper level. An instrumental course, with a stipulated handling of lots of data, according to Instructor 3, would be a good course to explicitly introduce these concepts in the first year of an undergraduate degree. Instructor 3 also pointed out that science teachers could perhaps first mention these Concepts of Evidence in high school, when students deal with various kinds of variables and sampling issues such as getting a representative sample of a population. Students in high school would have to determine whether the sample taken experimentally is really representative of the whole population. This sampling issue would also apply to other fields and disciplines such as social or genetic experiments, where only a segment of the population is sampled.

Instructor 6 was of the opinion Concepts of Evidence are being addressed stating “I’m sure we all have that in one form or the other.” All instructors have the goal of imparting knowledge to the students. To achieve this, one has to have a plan. The way the Concepts of Evidence are laid out is one such plan. Instructor 7 felt that these Concepts of Evidence are currently being taught in stages. The chemistry students are introduced to procedural knowledge “probably not as Concepts of Evidence, but as a normal part of learning chemistry.” The Concepts of Evidence are not taught as a separate part of the curriculum, but “they get that piecemeal, some application of that in chemistry.” In the general chemistry class they are introduced to Concepts of Evidence, and the instructors build on what the students learnt in high school. The instructors explicitly emphasize observation, actually acquiring one’s own data, and graphing it, but they are not introduced to some of the higher order, non-linear relationships. “Multiple measurements, error bars, instruments, calibration and error” are only introduced in second-year analytical chemistry courses.” Instructor 7 stressed that in analytical chemistry:

> I think we do do some of it (i.e., Concepts of Evidence) in the second year in the quantitative analysis course. That, of all our courses, is very, very lab-driven. Students definitely don’t design their own experiments there, but they do have to make judgments about their own work, and they have to interpret their own data. And their grade hangs on whether they make the right decision or not. (p. 14)

Therefore, according to Instructor 7, you find students stressing over the lab results for the first time, and actually worrying about the results that they are getting!

As far as designing their own experiments, Instructor 7 stated students did not do a lot of that at the undergraduate level, until they have reached the fourth year. Several chemistry undergraduates, in their fourth year, do “two terms of investigative research,” which entails working one-on-one with a professor or a post-doc or graduate student. Such a research project “... brings in the entire level of concepts. The undergraduate student who finishes that, possibly with a paper and lot of drive to publish, has been through some of the higher order things” (i.e., Concepts of Evidence). This means that some Concepts of Evidence will be used only in the undergraduate fourth-year research project. Thus Instructor 7 has clearly shown how the teaching and learning of Concepts of Evidence is relevant from the first-year to the fourth-year of the chemistry undergraduate program.

Instructor 8 was of the opinion that Concepts of Evidence are a part of chemistry. They are little pieces of information introduced here and there. When discussing phenomena such as error bars, calibration of equipment and zero point error, these terms are used all the time in chemistry, but probably not labelled, “this is a Concept of Evidence.” Except for the label “Concepts of Evidence,” these constructs are applicable in chemistry, and promote the constructivist pedagogy.

I was happy to learn that most instructors thought that Concepts of Evidence are very applicable in chemistry and that they support the constructivist pedagogy. The majority of instructors felt that it was very relevant and necessary to teach these Concepts of Evidence at the undergraduate level, beginning in the first year and continuing till the fourth year when some of the more complex concepts linked to the design of experiments could be implemented in the students’ investigative research projects. Duschl, Schweingruber, and Shouse (2006), stressed how important instruction in the classroom was on the learning of how to control variables, and achieve the conditions required for a fair test. One instructor, in my study, was particularly insistent that the teaching of Concepts of Evidence should be confined to the undergraduate level and should not be commenced in high school.

The second finding: Teaching more effectively. The second finding was centred on whether the teaching of chemistry can be made more effective by introducing these Concepts of Evidence. A third of the respondents replied strongly in the affirmative, with comments such as "Certainly, I think that stuff, talking about something like the Concepts of Evidence, this is a good thing because you are going to ask students to think about ... ." Some instructors, however, had reservations about the number of these concepts which could be addressed in class due to a tightly-packed curriculum resulting in a shortage of time. One respondent (i.e., Instructor 4) replied on the spur of the moment, "My first gut instinct would be to say "no." Instructor 4 then added:

> I don't see either a great deal of value come from explicit, explicitly telling them, "These are the points you need to know, these are the points you need to think about or these are the points we are looking for." I find it more beneficial for them to discover these things on their own. (p. 21)

Three of the answers from the respondents fell into the "Other" category. The individual contributions of these three instructors are explored below.

Instructor 4 gave two answers to the question "Can the teaching of chemistry be made more effective by introducing these Concepts of Evidence?" The first answer, quoted above, was in the negative, but later changed to the second answer, "It probably wouldn't hurt." Instructor 6 reiterated Gott et al.'s (2003) way of laying out the Concepts of Evidence: "I think this sort of lays out in detail what ought to be done in order to make it successful. And I think it's a good layout for them." When asked the same question as above, Instructor 8 elaborated by saying:

> Would it help? Maybe. Would it hurt? No. ... Maybe in the sense of if we had a whatever "x" number of minutes of lectures devoted to saying, "here are Concepts of Evidence, here are details which we expect." Like, we have to worry about zero point

> error, we have to understand fitting, and as the term progressed then you incorporated some ...So would it hurt to mention this at the beginning of your classes? No, it wouldn't hurt. (p. 12)

Instructor 8 did emphasize that it would not be possible to include all the Concepts of Evidence, but may be a few could be introduced. This instructor is positive that including the Concepts of Evidence into the first-year teaching of chemistry would have no adverse effects, but had reservations about the number of Concepts of Evidence and the time involved. As for the teaching of chemistry being made more effective by introducing these Concepts of Evidence, Instructor 8 commented:

> More effective? I don't know. The jury is out on that for me. I don't know if I can make it more effective. I think, I mean, you can make people think about this and they may appreciate it and get a better understanding for the science they do, but for me teaching chemistry, I don't know. I'm not too certain at the moment. (p. 14)

Although Instructor 8 did acknowledge that introducing Concepts of Evidence may increase the students' appreciation and understanding of science, this instructor was not convinced that the teaching of chemistry could be made more effective in this way.

If Concepts of Evidence are taught piecemeal from the first year, by the time these undergraduates doing science courses get to the second-year analytical chemistry courses, they have the extant knowledge required to build their new understandings of procedural knowledge. This is consistent with the constructivist pedagogy which takes into account what students know and scaffolds the teaching to build on that initial knowledge. In this manner, by utilizing the constructivist pedagogy in the teaching of Concepts of Evidence, teachers can "build knowledge structures that are commensurate with knowledge of the discipline" (Tobin, 1993, p. 7).

Theme Four. Teaching Concepts of Evidence: Instructors' Views. One of the themes that emerged from the interviews of my instructors was their views about the teaching of these

Concepts of Evidence. Instructors sometimes had opposing views about the pedagogy of these concepts which are the building-blocks of procedural knowledge.

The first finding: Two schools of thought about teaching Concepts of Evidence. The most interesting issue that arose was that of having two schools of thought about teaching Concepts of Evidence. One school of thought in the minority was that chemists acquire the necessary procedural knowledge by working on projects and ideas, struggling with them, and dealing with complicated issues. "They don't get it by your standing there and telling them, 'These are the Concepts of Evidence,'" said Instructor 4. According to this instructor, a better way of teaching these Concepts of Evidence would be to keep them "below the surface," and allow the students to explore them for themselves. Instructor 4 was convinced that was the way to make the students really learn these Concepts of Evidence, and further added:

> ... I think it's better to teach those concepts in a "soft" way, as undertones to your education, not as explicit ideas. I think that's straight-jacketed. But maybe I'm just queer. That's not how chemists would work. I am not going to go through a list and say "Oh, look at all my Concepts of Evidence." You know what the evidence for a particular argument is. You know what errors are. You know how errors work. You get that through experience. (p. 25)

This instructor seemed to advocate that students struggling with these Concepts of Evidence work on their own, without any formal, explicit teaching of these constructs.

The second school of thought was that these Concepts of Evidence have to be addressed at some point and they have to be taught specifically. This was echoed by eight out of the nine instructors, who were all in total agreement and only wondered about when this teaching should occur. Most were of the opinion that these Concepts of Evidence should be taught specifically, as early as possible, beginning in the first year. Instructor 5 reiterated:

> ... I think it would be better if we emphasize more of that from the word "go." I can't imagine getting to fourth year and wanting to go to Grad School and not being able to collect data properly and so I think it needs to be started in first year. (p. 4)

Then, Instructor 3 brought up an extra point that it would be difficult to learn all this material on Concepts of Evidence in one year. The researcher replied that building on the students' knowledge of these Concepts of Evidence over the years may be a better idea. Instructor 3 agreed wholeheartedly and added, "It's almost like taking baby steps, ... before you can run ... you have to walk."

Instructor 7 expressed the view that "even in our current structure," Concepts of Evidence are covered "as a normal part of learning chemistry." The chemistry student is taught these Concepts of Evidence "piecemeal." Instructor 7 went on to add that most students at that institution take a first-year statistics class, during which they amass a lot of the ideas that they need to know about data manipulation and treatment. Teaching about Concepts of Evidence should enable students "to think on their feet, ... how it was designed, and how it could be designed better or why there are certain things that are done."

In addition, Instructor 7 expressed uncertainty as to whether teaching a unit on just Concepts of Evidence alone would be beneficial for the students. A lot of content that is taught is curriculum-driven, and the curriculum itself is so "packed" that there is no time to include the Concepts of Evidence. Bringing Concepts of Evidence into the lab is also problematic because there are large numbers of students, leading to the formation of many sections. How does one grade all these sections consistently? How do they all get good guidance in learning how to design an experiment? The teaching assistants also have different strengths. One teaching assistant may be very happy to teach labs and may be good at it, while another may not be too keen about it. Consequently, achieving consistency is very hard in the large first-year classes.

The second finding: Level to teach the Concepts of Evidence. When asked at what level the teaching of these Concepts of Evidence need to be pitched, Instructor 2 had a very interesting comment, insisting that the teaching of Concepts of Evidence should be at the first-year undergraduate level and not in high school:

> I have a problem with high school trying to get their paws into all sorts of things that are high level, and they don't get the basics. I think that high school needs to go back to teaching the basics, and let postsecondary people do, you know, the higher level things. High schools can teach the basics, and show the applications so that they understand the societal applications of what they're learning. But to go into something like this, this is not a place for high school. I'd rather they learnt their chemistry a little better rather than be able to have this kind of critical thinking at this level. They don't have enough body of knowledge to do any practical work with this kind of work. (p. 7)

According to Instructor 2, the high school teaching should be limited to the basics. Higher level thinking tasks, such as critical thinking skills and students judging their work and that of others, should be left to the postsecondary institutions, and should not even be attempted at the high school level. This instructor is absolutely certain that the explicit teaching of Concepts of Evidence should be commenced only at the first-year undergraduate level.

This educator expects high school teachers to equip their students with a body of knowledge and to procure in them an ability to do math. However, in high school the curriculum developers have decreased the body of knowledge taught. Therefore, the postsecondary instructors have to provide their students with the knowledge they missed in high school. Instructor 2 was quick to point out:

> They know how to work in groups. They know how to do all that kind of stuff, but the body of knowledge, they don't have it in them. They can look it up in a data booklet or on the computer. But they don't actually know anything. (p. 8)

Instructor 2 expected students leaving high school to have a solid and fundamentally basic repertoire of knowledge to fall back on when faced with decisions in chemistry classes or labs; this, according to Instructor 2, seems to be lacking from students graduating from high schools today.

Theme Five: Improving the research activity. The first two findings under this theme were suggestions offered by the instructors during their interviews, on how to improve this research activity. The methods used to collect and represent evidence comprised the latter findings of theme five. I purposely left this theme open and focussed on the way evidence was *collected* and *represented* in the APL itself. In addition to being the medium used to test the understanding of Concepts of Evidence held by first-year undergraduate students in my study, the APL had the extra function of bridging the gap between the language of science and the language of school science (Phillips & Norris, 2009). In the research laboratory, scientists are attempting to create new knowledge, while school science in the classroom is trying to develop students' understanding of old, established scientific knowledge (Osborne, 2009). The two aforementioned contexts, therefore, have fundamentally different goals. Furthermore, it is important that the language related to evidence is clear and transparent, shedding light, where possible, on the line of reasoning employed by the authors. The third and fourth findings under this theme are instances where these lines of reasoning were not transparent. Some Instructors picked up and commented on these cases where the reasons behind the collection and representation of evidence were not explained or accounted for in the APL or even in the primary literature. The fifth finding, therefore, emerged as concerns some instructors expressed about the reliability and validity of the research activity described in the APL.

The first finding: Choosing primary literature to write an APL. In my study I wanted to choose a peer-reviewed journal article that had some application and meaning for the general student population at the first-year undergraduate level. The chemistry professor in whose class I piloted my evidence survey and the interviews, gave me a few papers that he and his research group had published in peer-reviewed journals. Out of the ones he gave me I chose the one published in the Journal of Photochemistry and Photobiology titled "Molecular Beacon Probes of Photodamage in Thymine and Uracil Oligonucleotides" (Yarasi, McConachie, & Loppnow, 2005). The main reason for my choosing this paper, as discussed in Chapter Three, was that I thought students would find it interesting because it has the practical application of demonstrating how uv-light induced damage can be detected in DNA or RNA using molecular beacon probes which emit striking fluorescence intensity when these nucleotides are damaged or mutated. In addition, I wanted to choose an article which could be understood by first-year undergraduate students, and I felt this paper fulfilled these conditions.

During my interviews of the instructors, I asked for their opinion as to how to increase the response rate of the students. Instructor 3 voiced the opinion that the primary literature chosen should be selected carefully to incorporate topics which would appeal to first-year undergraduate students, and it must catch their attention just like an attractive advertisement. Instructor 3 added:

> Let's say that paper was on music, or your paper is on Viagra, suppose you throw something out like that, you can bring in a lot more people. ... It's almost as though you've got to advertise. Sex sells, right? So we are going to probe damage to the reproductive systems, or something to that effect, right? Even if it's acne. May be the acne would do it!
>
> (p. 14, 16)

Instructor 3 also made the point that since Chemistry 101 is a basic chemistry course the students may be looking for a paper whose content is related to topics covered in class.

Students were also asked during the interviews whether they enjoyed reading the APL. The majority (nearly three-quarters) enjoyed reading it, with about the same percent stating that the APL's most useful and exciting parts were the annotations and meanings provided in the right column. There were some, however, who found the initial reading of the APL a struggle and somewhat frustrating, but they were a minority. The experiences of this student minority, however, will be taken into consideration, in the future, when choosing a primary literature article to write an APL.

There was considerable mention in the literature of the idea of improving the student experience of frontier science by introducing primary literature in various ways into the science classroom. Pall (2000) outlined the main features to bear in mind when choosing peer-reviewed primary literature for a general science education course aimed at non-science majors: (1) There are many topics of “great intrinsic interest to students” (p. 258); (2) A lot of primary literature (especially the medically-oriented journal articles) is readily accessible; (3) Students greatly value the opportunity to “search and analyze the medical literature” (p. 258), especially when it is applicable to their lives. All who might need to choose the primary literature to convert to APL – such as teachers, pre-service teachers, teacher educators or researchers in science education – would be well advised to bear in mind these three features mentioned by Pall.

The second finding: Improving the response rate. As the reasons for the low response rate were explored, Instructor 3 suggested that tying this whole research activity to marks might be a way of increasing the response rate, since the students are very mark- and- time-oriented. “I would say, you know, 3% of the grade, is probably worth like a $100.00 for them or more than that.”

Another suggestion that Instructor 3 had for improving the response rate is to select a paper that caters more to what the first-year undergraduate students know as chemistry. The cautionary note provided by the researcher was that one had to choose a chemistry paper in which all the main Concepts of Evidence were embedded. That is a Herculean task, and one that cannot be underestimated. Instructor 3 agreed and added:

> It starts to get tougher, and certainly when you know, let’s say someone is actually interested in participating, right? They might start off, and once they start reading, they’ll

> go ... eh ... no ... you know, I'm not finding this very familiar or comforting to what I know. It's their background, right? And, if you have a paper that you know, if it's more related to their chemistry ... (p. 15)

When asked what chemistry would be most appealing to the students, Instructor 3 complimented the researcher, saying that was a good question. Instructor 3 was of the opinion that any paper that monitored bodily functions or malfunctions, such as diabetes, would appeal to the students, and added:

> Something that will really make them sit back and go "Oh yeah! I think I'm going to learn something here, you know about .... just my surroundings. There's always going to be also, like ... it could be a cultural thing, right? Some people are just going to be turned off by certain papers. So, I don't know. It's a tough one, for sure. (p. 16)

In some ways, one has to choose an article that sells and appeals to the students. There must, however, be points in the paper where Concepts of Evidence are explicitly expressed or implicitly implied. As a researcher, when choosing a primary literature to adapt, it can be difficult to balance the desire to appeal to the students with the duty to teach the Concepts of Evidence. Instructor 3 also voiced the opinion that having the pre-test and post-test done in class, instead of online might encourage more students to participate in and complete the tasks.

The third finding: Collection of evidence in the primary literature. The third finding probed the reason for collecting evidence within the 25-60^0C range in the experiments done by Yarasi et al. (2005). I found that a third of the instructors believed that this limitation was employed to maintain the collection of evidence (i.e., the observation of the fluorescence intensity) within the physiological temperature range. Approximately 10% of the respondents were unsure of the answer, while half of the instructors chose the "Other" category offering a variety of answers. Instructor 1, for example, explained that the restriction was imposed "... because most organisms live between 25-60^0C." Instructor 2 offered three reasons for this restriction: first, it could be this "thing" decomposes at about 60^0C, and therefore this trend is no longer seen; second, this could be the temperature range of interest for what these authors are studying; and third, this could be due to the limitations of the instruments used for measurement. Instructors 3 and 5 offered the explanation that the degradation or denaturation of the molecular beacons occurred at temperatures outside the range studied. The most lucid answer was offered by Instructor 7, who reasoned that "... if you look at the melting points of these things, they are somewhere around 30^0C." This is actually inaccurate, because on examining Figure 2 in the APL, it is evident that the melting point is actually higher - halfway between the hair-pin structure and the random coil structure - around 45^0C. Further, Instructor 7 argued that room temperature is probably between 20-30^0C, and that is why Yarasi et al. (2005) started measurement at that temperature.

I learnt from the authors that they chose the temperature for the melting point as 45^0C and placed that as the centre. Then they took enough points on either side of the range dictated by their solvent, water. When they use water, they cannot go much below 5^0C, and not much above 85-90^0C. The authors then selected as large a range as possible within those limits (i.e., 25-60^0C) that is necessary to observe both the hair-pin and random coil structures, and their fluorescence intensity levels.

In the primary literature (from which the APL was adapted) for my study, however, I was surprised that there was no explanation of why evidence was gathered in the 25^0 to 60^0C range by the authors (Yarasi et al., 2005). I explained these omissions, which were picked up by some participants, by emphasizing how this primary literature was an article in a peer-reviewed journal in the field of photobiology and photochemistry. Therefore, it must have met the requirements in these frontier disciplines of science. Besides, I reiterated the fact that this was the preliminary work done by this group of scientists. Further work done in this field by the same group has been more rigorous and has been accepted for publication by leading journals in these disciplines of science. These publications, however, were not available for use in my research activity because at that time they were still “in press."

The fourth finding: Representation of evidence in the APL. The fourth finding under this theme dealt with how evidence is represented in the APL and examined the kind of information that is not provided by the graphs in this document. Just under half of the instructors correctly identified error bars as information that is not provided by graphs in the APL. They added that the reason for this omission was that the error bars were quite small - as small as the points on the graph. Half of the instructors’ answers, however, fell into the “Other” category. Instructor 1 insisted that “...what the authors are telling you is that they’ve measured this more than once and ... they didn’t see significant differences.” This is a possible explanation, but there is no proof in the paper by Yarasi et al. (2005) that these experiments were repeated multiple times, an opinion that was echoed as an omission by Instructor 6 as well. Instructor 7 suggested that the authors may have chosen to show one result over an average result, adding that if the error bars were large, it would be difficult to put them on one plot anyway. Instructor 8 displayed excellent reasoning, by providing three reasons why error bars were not included. First, error bars are small, smaller than the symbols representing a point on the graph. Secondly, error bars were used for fitting purposes, to decide which points are real and which have to be modified to fit the curve or line obtained in the graph. Thirdly, this is not an entirely quantitative study and, therefore, error bars are not significant and can be neglected. Here, the authors are trying to demonstrate a trend rather than absolutely fitting the data points into a line or curve. This was a well-reasoned argument that was very convincing and found to be true.

The fifth finding: Reliability and validity in the primary literature. The fifth finding was the opinion that instructors had of the reliability and validity of the primary literature (i.e. the whole investigation described in the APL). During the interviews I sought to find out whether the instructors thought the reliability and validity of this study was sufficient, or whether they found both a little lacking in this investigation. Some of the instructors were of the opinion that the level of reliability or validity in the paper is sufficient. They claimed that this is preliminary work, but the results obtained since then have been consistent - especially the structure of molecular beacons and how they work. This proves that the molecular beacons are quite reliable in detecting DNA damage. 20% of the participants declared that the level of reliability or validity was sufficient. Since this work has been published in a peer-reviewed journal, the instructors assumed that other people in this field had deemed it sufficiently reliable and valid. A few of the instructors, however, were unsure of the answer.

The answer provided by Instructor 2 consisted of three main points. The first point was that heating and then cooling are a means of triangulation – two ways of testing the same thing. The second point was that it was uncertain whether these researchers really know if this fluorescence intensity actually measures the damage. In other words, is it really measuring what it claims to be measuring? The third point that Instructor 2 raised was that there is no evidence in this diagram for reliability of data, because it is not clear whether the points shown in the figures are aggregates or single points.

Instructor 3 had a number of concerns about the reliability and validity of this experiment. This instructor felt that the fluorescence intensity was a little worrisome. Since the article is in a peer-reviewed journal, Instructor 3 was not sure whether the APL was "just conforming to the standards of this area of photochemistry and photobiology." This instructor was also concerned that the data did not show error bars. The way the data were presented in the figures gave the impression that the results shown were from just one experiment, whereas that might not even be the case. Instructor 3 added:

> So I heated it up. Here are my points, and then I cooled it down. Here are my points. I haven't really shown you that I have done more than one, right? And you know the fact is ... I'm sure that it probably has been done more than once, but as you say, it is not really lending itself in telling us that. (p. 13)

Instructor 3 maintained that the authors should have indicated very clearly in the paper how many trials had been done. This could have been done using error bars or some other means of conveying clearly whether there were multiple trials or only a single attempt.

Instructor 5 was first of the opinion that perhaps the trials had been done a number of times, but after examining the figures in more detail, there was a change of mind. Then this instructor stated:

> They don't talk ... like I guess that kind of puts in some ... you wonder how scientific it is when it says arbitrary units ... So, you could have put the units in and put it in as a Δ, in the end. You could have put that in. They didn't say, you know, we did 20 trials and that is the average of them. So, ... (p. 3)

I agreed that there was no indication in the paper as to how many times the trials were done. However, I pointed out that this paper was just the preliminary work in this area:

> EV: Now I understand that they have done some further work on this. So it is valid, but it's just the way it has been presented.
> I5: It is a little ... a little bit "iffy," but ... (p. 3)

So far in this chapter I have described the findings and the themes that emerged from the analysis of the interview data. I conclude the findings chapter with an analysis of the data collected from the evidence survey.

Data from the Evidence Survey: A research Preactivity

The marks that the 61 students who participated in the evidence survey achieved on the pre-test and post-test were analysed. The post-test mark was higher than the pre-test mark for 53 students out of the total of 61. This means that the majority (86%) of students had a higher score on the post-test than on the pre-test. The increased learning is due to the pedagogic treatment which is the teaching intervention - the only research activity placed between the pre-test and the post-test. During the teaching intervention the initial, extant knowledge the students possess at the pre-test is connected to new understandings developed as the students go through the experience of watching the video (i.e., the teaching intervention) and answering the post-test. The proof provided here, therefore, is that Concepts of Evidence can be taught and learnt. This suggests that the answer to the initial question of my research project, "Can Concepts of Evidence be taught or learnt?" is in the affirmative, because the majority of the student participants performed better on the post-test than the pre-test.

In order to effectively answer my second research question, which was about the difference the teaching intervention made to the initial understanding that students had of the Concepts of Evidence and of the substantive knowledge in the APL, I had to tease out what led to the differences that I saw between the students' performance on the pre-test and the post-test. The values of the differences between the post-test mark and the pre-test mark (i.e. post-test mark minus pre-test mark, known in this study as Δ_{diff}) are not very stable. This analysis had to be made more valid by expressing the post-test as a covariate of the pre-test which is the baseline (as discussed in Chapter Three).

According to the probability theory, covariance is a measure of how changes in one variable are associated with corresponding changes in a second variable. Covariance also shows the extent to which two variables are associated linearly. In other words, covariance demonstrates the joint variability of two random variables. If greater values of one variable correspond with greater values of the other variable, the covariance is positive. Similarly, if lesser values of one variable result in lesser values of the second variable, the covariance again has a positive value. For example, if a balloon is blown up, it gets larger in all dimensions. Conversely, if greater values of one variable mainly correspond to the lesser values of the other, (i.e. if the variables show opposite behaviour) then the covariance is negative. For example, if a sealed balloon is squashed in one dimension, then it will expand in the other two dimensions. Covariance or the sign of the covariance therefore, shows the tendency in the linear relationship between the variables.

The two variables considered here are the pre-test mark and the post-test mark obtained by each student participant. The covariance here is expressed as the covariance ratio that is obtained by dividing the post-test mark by the pre-test mark which is the baseline. This covariance ratio (i.e. post-test/pre-test) was calculated for all students who participated in the evidence survey. The calculated covariance values were used to divide the students into three groups: the first group had high positive covariance ratios above one; the second group had lower, average covariance ratios which were also above one; while the third group had covariance ratios below one. In the first group the post-test marks attained by the students was much higher than the pre-test mark. Therefore, this was the group where the teaching intervention was most effective. In the second group the post-test mark was still higher than the pre-test mark but the margin between them was much narrower than in the first group. In other words, the teaching intervention raised the mark the students obtained on the post-test in this group, but not by as great a difference as in group one. These two groups together included fifty two out of the total of sixty one student participants (i.e. 86%). The students in the third group, on the other hand, fared better in the pre-test than in the post-test. This was the group, therefore, where the teaching intervention was least effective. Nine out of the total of sixty-one students (i.e. 14%) who participated in this research activity fell into this category. A few student responses were chosen from each of the three groups above and their interview transcripts were analyzed qualitatively using content analysis.

In order to tease out the differences that I saw between the pre-test and the post-test I had to consider: (1) How did the students' learning styles contribute to the pre-test scores? ; and (2) How did the professor's teaching methods portrayed in the video contribute to the post-test scores? The students' learning methods were investigated by taking into consideration their pre-test scores as well as qualitatively analyzing their impressions of and attitude towards the APL. The professor's teaching methods, on the other hand, were probed by the students' scores on the post-test, as well as a qualitative analysis of their impressions of the teaching intervention and its delivery. The interview transcripts of the chosen students in each of the three groups were analyzed qualitatively, focussing especially on their answers to the question which asked the students about the APL and the question which interrogated them about the teaching intervention.

Student 45 (S45) who was in Group 1, stated that the APL was quite boring because this student's passion was geology. This student scored poorly in the pre-test securing only about 20% which was a failing grade. Student 45's interest and innate ability to learn were not well-aligned to tackle the questions of the pre-test effectively. However, S45 scored 50% for the post-test. He insisted that since he had read the APL prior to listening to the Instructor in the video, "that reinforced the material" (p. 28). The teaching intervention served such a validating function for many students especially those in Group 1. The students examine their own understanding while listening to the video: this examination validates what is right in their own understanding and highlights what is wrong. S45 whose post-test showed a marked improvement reiterated, "So he wrote on the board, explaining it and drawing diagrams whereas when you are reading it, you just read it and you see a graph, not actually drawn or explained" (p. 28). In the video, the professor said "Then the fluorescence would undergo an exponential decay" and immediately drew in on the board, as an afterthought, what an exponential decay would look like on a graph. This was doubly reinforcing and powerful because the students learnt when an exponential decay occurred, what it was called and what shape it took when represented graphically.

Student 53 in Group 2, found the APL "kind of dry" (p. 47) but "... really enjoyed the on-line component" of the teaching intervention" (p. 48), stating "I guess it is a nice change of pace from having to just read all the time, you can watch the lectures" (p. 50). This student also appreciated having the teaching intervention as an on-line video, because it could be accessed at any time and any part of it could be watched as required. This was a huge asset according to S53, which enabled this student not only to face the post-test with more confidence but to score higher on it. Student 40 in the same Group 2, found the APL "a little difficult to read. It wasn't something that I had studied before ... I had done RNA and DNA structure but not this in-depth" (p. 38). When asked about teaching intervention, S40 rejoined "Oh, it was really helpful. I thought it was excellent ... you could see how he labelled everything and kind of hearing him say it, and seeing the picture at the same time, kind of made things connect a lot better" (p. 42). Again this student seemed to be implying that the double sensory stimulation of drawing while speaking was important. From this student's perspective, seeing the picture at the same time as hearing what was said in explanation, enabled similar ideas to interconnect and make more sense. In the same vein, Student 29 when asked whether the better performance in the post-test was because this student had seen similar questions in the pre-test, replied emphatically "I shouldn't think so. But there was a pretty big difference in how much I understood the paper better after watching the teaching intervention" (p. 37). This was clearly reflected in S29's score of 14.75 marks in the post-test when the average was 11.4 marks. The teaching methods employed by the professor in the video seemed to have contributed immensely towards S29's performance on the post-test.

When Student 51 (S51), in Group 3 was asked whether he would have understood the primary literature replied, "Not completely, but to read this (pointing to the primary literature), and then to be given this (i.e. the APL), I think would have been more useful because then I would have to struggle through this a bit and have to you know maybe look it up on my own, and maybe do a little bit of extra research to understand it more" (p. 31). S51 seemed prepared and willing to take on the challenge of reading the primary literature and grappling with it because according to him "What I found was that I knew a lot of this stuff, so it wasn't extremely difficult for me" (p.31). On asking him whether he enjoyed reading the APL, S51 was quick to add, "I thought it was interesting. I didn't know how the molecular beacons have a Light and Dark Phase. That was interesting" (p. 31). S51 demonstrates a good understanding of the APL, foremost because of the mark attained on the pre-test (mark of 10.5 when the average mark was 9.4), and also because of the obvious ability to appreciate the finer, deeper details of the structure of molecular beacons. When questioned about the efficacy of the teaching intervention S51 responded: it "put pictures to things that I might not have thought visually in the head from reading the article ... I think the teaching intervention helped to a point" (p. 34). It seemed that S51 had some reservations about the efficacy of the APL, which were not quite pronounced here. Therefore one can only speculate what those reservations were. I further asked S51 whether the teaching intervention just boosted confidence or really helped this student understand the APL.

This student insisted “No, I think it really helped to improve my understanding” (p. 34). However, it must be noted here that S51, like all members of Group 3, performed better on the pre-test than on the post-test. This could be because of S51’s inherent interest in the content knowledge of the APL, and this student’s determination to “grapple” with the substantive ideas until he understands them. Therefore, more explicitly, here is how I have responded to research questions one, two, three and four.

Summary

My data have two components: the student interviews and the instructor interviews. Using thematic content analysis, four themes were identified from the student interviews and five themes emerged from interviewing the instructors. The findings that contributed to each theme were identified, and the manner in which each theme was construed, was traced elaborately.

The evidence survey demonstrated clearly that the majority (86%) of students performed better on the post-test than on the pre-test. This answered affirmatively my initial research question of whether Concepts of Evidence can be taught or learnt. Because of the controls implemented in the design of the evidence survey, the increased students’ understanding displayed in the post-test could mainly be ascribed to the pedagogic treatment which was the teaching intervention.

The efficacy of the teaching intervention was unequivocally established because it was the only research activity placed between the pre-test and the post-test. As such it played a substantial role in furthering the understanding of students. The dual stimulation of auditory instruction and visual imagery (i.e. drawings and flow-charts) increased the students’ ability to understand and retain what they were taught in a “conversational,” teacher-centred style through the video. The teaching intervention was most appreciated by the students for being well-paced with lucid explanations, very illustrative diagrams and anecdotal accounts of what it was like to carry out the initial experiments described in the APL. The instructors interviewed also had the opinion that the teaching intervention had tremendous efficacy in furthering the students' understanding of both substantive knowledge of the APL and procedural knowledge (i.e. the Concepts of Evidence embedded in the APL).

The interdisciplinary nature of the APL posed challenges for the students in understanding the substantive knowledge it entailed. For example, initially many students did not perceive molecular beacons as probes for detecting photodamage in DNA/RNA. Some areas of the content knowledge in the APL were challenging for the instructors as well, such as how the underlying structure of the molecular beacons determined their heating and cooling curves.

The students had to mentally wrestle with Concepts of Evidence to which they were just introduced during the teaching intervention. They may have had erroneous perceptions of the fair test, for example, which they recognized during the pre-test and clarified as they watched the teaching intervention.

Contrary to traditional views, it has been recently shown (Roberts et al., 2010) that teaching substantive knowledge on its own was insufficient, but has to be supplemented with the teaching of the underlying procedural ideas as well. Most instructors interviewed in my study thought that Concepts of Evidence are very applicable in chemistry and that they support the constructivist pedagogy. The majority of instructors supported the idea that it was very relevant and necessary to teach these Concepts of Evidence at the undergraduate level, beginning in the first year and continuing till the fourth year when some of the more complex concepts linked to the design of experiments could be implemented in the students' research projects.

If Concepts of Evidence are taught “piecemeal” from the first year to these undergraduate students taking chemistry courses, by the time they get to the second-year analytical chemistry courses, they have the extant knowledge required to build their own new understandings of procedural knowledge. This is consistent with the constructivist pedagogy which explores what students know and scaffolds the teaching to build on that initial knowledge.

However, there were two schools of thought among the instructors about the teaching of Concepts of Evidence. The majority of instructors thought that all the Concepts of Evidence have to be addressed at some point and they have to be taught specifically, beginning in the first year. A minority of instructors, however, felt that students should struggle with these Concepts of Evidence, on their own without any formal or explicit teaching of these constructs.

CHAPTER FIVE

Prelude: The Harvesting

Harvest time has arrived at the golden-coloured paddy fields in the Indian town of Aluva. The yellowish-brown rice husks are ripe enough to be harvested. The paddy fields have to be drained before cutting the ripened rice plants. Harvest time draws women to the paddy fields, because they do most of the harvesting. Quite early in the morning you can see women heading towards the paddy fields carrying their lunch box and their sickle. Armed with sharp sickles they plunge into their allotted segments. They cut the rice stalks as close to the soil as possible, and lay them in their segments. The men walk around tying the cut rice plants into bundles which are then carried on their heads to the winnowing site.

There is an air of expectation in the air. The farmers – both women and men – are brimming with expectation to know what kind of a crop they had this year. Before this can be known, the farmers have to separate out the ripe grains and remove the husk. This is very similar to what I will be doing in Chapter Five, where I will delve, with great anticipation and interest, into the meaning of my findings.

Before the quality and the quantity of the rice grain can be established, the harvested rice has to be threshed and winnowed. Threshing is the process of loosening the edible part of the rice grain from the scaly, inedible husk that surrounds it. This is mostly done by the women workers. The common method of threshing by hand is to separate the grain from the inflorescence by the impact of hitting the panicle on a large stone or log. The rice grains are threshed (beaten) in order to separate them from the stalks. Then the crop is spread over a mat or canvas, and women workers, holding on to a strong pole, trample the crop with their own feet. After the trampling, the straw is separated from the grain.

Winnowing is used to clean the grain and to separate out the rice from the husk. Whole rice is put into a basket made out of jute called a *mooram* in the local language, and then thrown into the air. The wind is made to blow onto the whole rice as it falls. Consequently, the light husk is blown away while the heavy grain falls back into the *mooram*. This heavy grain is the end-product - the rice - which must be very dry before it is stored. The rice is spread out on jute mats, sheets of plastic, or concrete, to dry in the sun.

All this takes time and much concerted effort. So, too, my final chapter with the overview and conclusions of my study, took time and persistent effort to synthesize. The data had to be arranged to fit together like the pieces in a jig-saw puzzle, to paint a fuller picture of my study. My qualitative data from the interviews and from the evidence survey had to be "threshed" and "winnowed," to bring out the meaning of my findings and separate out the "husk," the irrelevant details and facts. The meanings derived from my findings answered my research questions and much more: they were the precious "rice" grains.

DISCUSSION OF FINDINGS

This chapter entails a discussion of the research findings and results furnished in Chapter Four. Where necessary, I introduce more relevant literature in this chapter. This is because, as Tashakkori and Teddlie (1998) pointed out, as a researcher, one has to be flexible, as it is difficult to predict at the start of the project which literature will be most relevant. In addition, such adaptability is vital for qualitative research (Marshall & Rossman, 2006), because that is what enables the work to be "played by ear: it must unfold, cascade, roll, (and) emerge" (Lincoln & Guba, 1985, p. 209). The discussion here will include an examination of: (a) An overview of the study: A response to the research questions; (b) The student experience through the research activity; (c) Conceptual Change; (d) Factors which influence teaching and learning; and (e) Conclusions.

An Overview of the Study – A Discussion of the Research Questions

My initial pre-research question was *"Can Concepts of Evidence be taught or learnt, and if so what are the factors which influence their teaching and learning?"* The findings from my qualitative data appear to answer this question resoundingly in the affirmative because the majority (86%) of the students who participated in my evidence survey showed an improvement in their performance on the post-test.

From the scores the students obtained on the pre-tests it was evident that they do not come to this research activity as *tabula rasa*, but had their own pre-instructional conceptions, which may be opposed to or in line with the scientific view. When these pre-instructional conceptions do not comply with the scientific view, they are referred to as misconceptions or alternate conceptions. The role of the teaching intervention is to address such misconceptions. It is hoped that if the teaching intervention is effective, students will be enabled not only to change their conceptions until they are correct, but to perform better on the post-test. In my study, since the majority of the students performed much better on the post-test, one can state emphatically that pedagogy does make a substantial difference in how students understand Concepts of Evidence. This understanding is not a genetically transmitted trait, but one which can be improved by using appropriate teaching interventions.

One set of research questions focussed on the teaching intervention, and inquired whether it made any difference to the student participants' understanding of Concepts of Evidence and substantive (content) knowledge. The majority of students praised the teaching intervention for its efficacy and usefulness, and were highly appreciative of the drawings and diagrams incorporated into the teaching. They felt that the auditory and visual stimulation helped them to better and more vividly understand the material being taught in both the substantive knowledge and procedural knowledge (i.e., Concepts of Evidence). Most of the student participants thought they had fared much better in the post-test because of the teaching intervention. The instructors were of the opinion that the teaching intervention had a significant role in improving students' understanding about the content knowledge as well as the procedural knowledge (i.e., Concepts of Evidence). Of the nine instructors interviewed, one predicted that the teaching intervention would have a tremendous effect on students' understanding.

Another research question raised was the view of instructors about including Concepts of Evidence when teaching chemistry to first-year undergraduate students. There were two schools of thought on this topic. The majority of instructors thought that it was a very good and feasible idea, and stipulated that teaching the Concepts of Evidence should be done "piecemeal" starting in first year, and continuing throughout the Bachelor of Science degree program right up to the final year. A few of these instructors emphasized that they were teaching some of these concepts in their undergraduate classes, but were not classifying them as “Concepts of Evidence.” One instructor, however, was of the opinion that students should not be taught these Concepts of the Evidence, but should be left to struggle with them on their own. The onus should be entirely on students to achieve this understanding.

Review of the Key Findings from the Interviews

From the data analysis in Chapter Four, I identified a total of nine themes, of which four were from the interviews of the students and five from interviewing instructors. The nine themes are summarized for this chapter in Table 1: each theme consisted of a number of findings.

The main pedagogical tool wedged between the pre-test and the post-test is the 32-minute video which is the teaching intervention. I evaluated the efficacy of the teaching intervention by the students’ overall impressions of the video, their assessment of the teaching intervention as a pedagogical tool, and finally by their expectation of how they had fared in the post-test. The students’ impression of how they had fared in the post-test was compared to their actual mark on the post-test.

Most instructors were convinced of the efficacy of the teaching intervention in promoting a good understanding of the Concepts of Evidence. The visuals and anecdotal explanations were lauded as being pedagogically sound in successfully linking empirical evidence to the underlying scientific theories. One instructor vouched that the teaching intervention would certainly further the students’ learning of the substantive knowledge and the procedural knowledge (i.e., Concepts of Evidence) contained in the Adapted Primary Literature (APL). Another instructor voiced the opinion that the primary literature chosen should be selected carefully to incorporate topics which would appeal to first-year undergraduate students. Yarden, Norris and Phillips (2015) confirmed that “choosing a suitable primary literature paper can be more challenging than the

Table 1

Summary of Themes and Literature in Agreement or Not In Agreement

Theme	Literature Supporting	Literature Not Supporting	New Contribution
From Students' Interviews			
1. Efficacy of the Teaching Intervention (TI)		(Zhou et al., 2005)	xxx
2. Grappling with the APL		(Norris et al., 2009b) (Yarden. 2009)	xxx
3. "Wrestling with CoEs: the views of students	(Gott et al., 2003) (Roberts et al., 2010)		xxx
4. Challenges for attaining Substantive Knowledge	(Janick-Buckner, 1997) (Herman, 1999) (Hoskins, Stevens & Nehm, 2007) (Falk, Brill & Yarden, 2008)		xxx
From Instructors' Interviews			
1. The Instructors' Views on the TI	(Phillips & Norris, 2009) (Osborne, 2009)		
2. Challenges in the Content Knowledge of the APL	(Yarasi et al., 2005)	(Cajete, 1999) (Aikenhead)	xxx
3. "The thinking gets kind of lost in the content"	(Roberts et al., 2010)		
4. Teaching CoEs: Instructors' views	(Roberts et al., 2010)		
5. Improving the research activity	(Pall, 2000)		xxx

adaptation process itself" (p. 7). In my study 65% of students really liked the APL while 4% found it boring. The process I engaged in to choose the primary literature has been described in detail in Chapter Two.

Pall (2000) outlined the main features to bear in mind when choosing peer-reviewed primary literature for a general science education course aimed at non-science majors: (1) There are many topics of “great intrinsic interest to students” (p. 258); (2) A lot of primary literature (especially the medically-oriented journal articles) is readily accessible; (3) Students greatly value the opportunity to “search and analyze the medical literature” (p. 258), especially when it is applicable to their lives. All who might need to choose the primary literature to convert to APL – such as teachers, pre-service teachers, teacher educators or researchers in science education – would be well advised to bear in mind these three features mentioned by Pall.

I did not specifically address the Concepts of Evidence in the APL and this was done intentionally because the focus of this document was to cover the primary literature. The Concepts of Evidence were actually embedded within this APL and were only addressed in the teaching intervention. Since Gott et al. (2003) classified these Concepts of Evidence into 21 categories with sub-categories under each, it was not possible to include them all in this research activity. Therefore, I chose just four categories of Concepts of Evidence which were embedded in the APL. Many of the Concepts of Evidence included were based on aspects of chemistry experimentation that first-year undergraduate students had met in their prior knowledge and experience of science either in high school or at the postsecondary level.

Half the student participants voted for the Concept of Evidence #12, “Design: Validity, Fair Tests and Controls” as the most important one. I learnt that the definition of a fair test ranged in accuracy from minimal to medium to full, and that a tenth of students had a fully accurate definition. Over half the students correctly identified the dependent variables and the independent variables in certain experiments performed in the APL. Out of the 54 students interviewed 41knew that only one independent variable was changed at one time. I was greatly encouraged by this result because it showed most students knew what a fair test was, and were aware of its importance. Since the Concepts of Evidence were embedded in the APL it was apparent to me how essential it was for the students to have a good understanding of the substantive knowledge, in order to comprehend how the evidence concepts in the research activity were applied.

The substantive knowledge in the primary literature was interdisciplinary in nature, pooling together content knowledge from the disciplines of photochemistry, photobiology, genetics, chemistry, biology, spectrophotometry, and fluorometry. When the substantive knowledge needed to understand this paper (i.e., Yarasi et al., 2005) is a combination of scientific disciplines, then the content becomes challenging especially for students who are not experts but novices to many of these fields (Newell, 2007). As the knowledge from the different disciplines is brought together in the primary literature, the various disciplinary contributions interact in a complex manner. This complexity, according to Newell (2007) "is embedded in the arguments developed to integrate that knowledge, the organizational structure employed, and in the writing process itself" (p. 108). It was this complexity which was simplified by writing the APL. The findings that constituted the greatest challenges for attaining substantive knowledge were the ideas from students on anomalous data and what to do when these occur in their own scientific investigations; experimental integrity – a serendipitous finding; an understanding of and recognizing "sampling"; setting up control experiments for specific reasons; and the overall purpose of the series of experiments discussed in the APL.

Findings from the Evidence Survey

The results from the analysis of my data from the evidence survey were dealt with in detail in Chapter Four. The overarching message from all my findings may be encapsulated as pedagogy does make a substantial difference in the learning of Concepts of Evidence by first-year undergraduate students taking science courses. This implies that the teaching intervention furthered the students' understanding of the Concepts of Evidence and the substantive knowledge contained in the APL. Other factors which may contribute to the improved understanding displayed in the post-test are the confidence level of the students when doing the tests, especially the post-test; the validity of the instrument used for testing the learning that has occurred; the human error; the instrumental error; the health of the student participants, especially their anxiety levels; and the time the students have to devote to my research project. Therefore, although there may be many factors which enhanced the improved performance in the post-test, the treatment effect is the major contributor.

Themes and Literature in Agreement or Not In Agreement

In most cases, the themes identified from my study were supported, to varying degrees, in the literature, as shown in Table 1. In a few instances, however, there was a lack of agreement between what other researchers have reported and what I have identified as findings in my study (Refer to Table 1). For example, Zhou et al. (2005) discussed only learning by the use of text and computer simulations and applets: there was no mention of a teaching intervention as adopted in my study.

Another example is where Norris et al. (2009b), and Yarden (2009) classify the genre used in APL as solely argumentative. The students in my study, however, classified the genre of the APL using two descriptors: (1) descriptive and argumentative; or (2) descriptive and narrative. Norris et al.'s (2009b) and Yarden's (2009) claims were contrary to what I learnt as the view of the students who classified the genre using two descriptors. Since both student trend classifications had “descriptive” in common, I understood that they perceived this as a common throughout the APL.

Experimental integrity was an interesting, serendipitous finding with ethical and moral implications that emerged out of the discussion on what one could do with anomalous data if one confronted them in one’s own laboratory work. The ideas which emanated from the student participants support the literature which informs us that when reading peer-reviewed journal articles on molecular biology and genetics, undergraduate students face highly technical jargon which is quite different from what they encounter in their own prescribed textbooks and popular newspaper reviews (Janick-Buckner, 1997; Hoskins, Stevens, & Nehm, 2007). This type of language-based unfamiliarity with substantive knowledge acts as a major challenge deterring students from reading primary literature (Herman, 1999), and even APL (Falk, Brill, & Yarden, 2008) in some instances.

The importance of substantive knowledge in comprehending the Concepts of Evidence and their usage were reiterated in recent literature. Roberts et al. (2010) claimed that “the role of procedural understanding is often neglected in science education” (p. 377). These authors took the stance that substantive and procedural knowledge are two distinct types of knowledge, though they can be connected. In addition, they claimed that it may even be necessary at times “to teach procedural ideas separately from substantive content” (p. 377). In the teaching intervention employed in my study, the chemistry professor who presented first taught the substantive knowledge to understand the APL and then turned to teaching four Concepts of Evidence (i.e., procedural knowledge) embedded within the APL. Since the teaching intervention was where the majority of participants were first introduced to these Concepts of Evidence, I realized that some students cognitively wrestled with some of these familiar concepts to perceive how they were central to evaluating evidence.

The studies cited in the literature have used a wide variety of conceptual change teaching approaches such as: (1) laboratory work (Akkaya, 2003); (2) conceptual change text accompanied with analogy (Bozkoyun, 2004); (3) an argument approach leading to an argumentation model (Zhou, 2010); (4) computer-based applets (Zhou, Brouwer, & Nocente, 2005); (5) small group discussions and hands-on activities (Van Driel, 2002); and (6) computer aided instruction (Tezcan & Yilmaz, 2003). However, teaching intervention in the form of a video sent to the participants on-line is a new approach that has not been mentioned as yet in the literature. This is therefore a new contribution made by my study to the literature in this field.

In the case of some of the other themes, I have described findings which were not dealt with (at least not dealt with directly) by other research. These were my new contributions to a particular theme, and to specific findings under that theme, as indicated in Table 1. On further analysis of the key themes that emerged by interviewing the students and the instructors, I realized that the key themes in each category are related to each other and fit well together as illustrated and explained below.

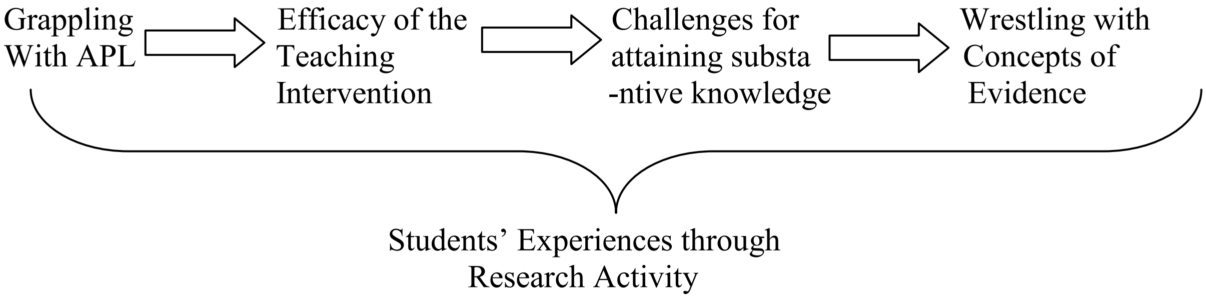
Conceptual Framework Showing Relationship between Key Themes

For Students

As a summary, I have provided Figure 5 showing the conceptual framework indicating how the key themes, identified from the student interviews, fit together and relate to one another. The challenges the students face to understand the substantive knowledge in the APL, are ameliorated by the teaching intervention, but they still have to struggle with understanding the Concepts of Evidence within the context of the APL. All these processes together constitute the

Figure 5

Conceptual Framework Showing Relationship between Key Themes from Students



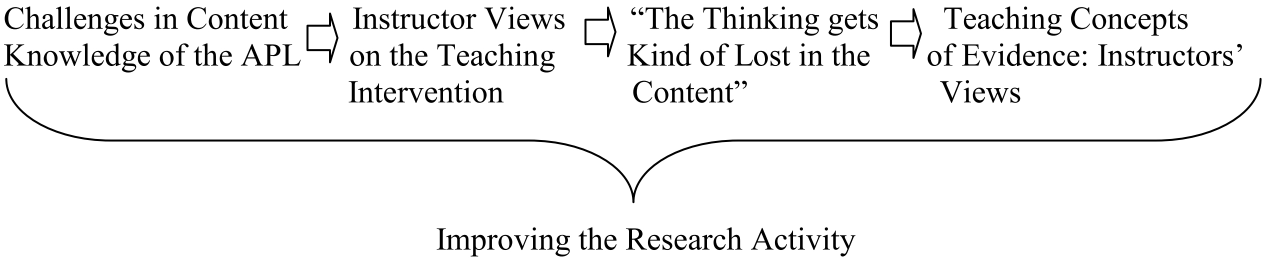
experiences of the students as they proceed through the whole research activity.

For Instructors.

Similarly, Figure 6 shows the relationship between the key themes obtained from the content analysis of the qualitative data collected by interviewing the instructors. The instructors as well were challenged by the substantive knowledge of the APL. However, because of their deeper scientific knowledge and greater experience with reading primary literature, they were able to identify the strong points and the omissions in the way evidence was represented in the APL. At times they felt the ideas of the Concepts of Evidence could be learned only by getting immersed and attaining a deep understanding of the content knowledge. They also offered some suggestions on how to improve this research activity.

Figure 6

Conceptual Framework Showing Relationship between Key Themes from Instructors



The Student Experience through the Research Activity

I contemplated the experiences of the students in reading the APL and going through the whole evidence survey stage by stage. The negative aspects such as the noise due to the movements of the camera, failing to fully understand the APL, the survey taking longer than expected, and the time of year at which it occurred, all informed me what it was like for the students to go through the survey. I learnt from my participants that these negative aspects of the evidence survey did not deter the students from appreciating its positive aspects which they found to be more numerous, more substantial, and more valuable.

From their description of the positive aspects of this research activity, I learnt that the students valued certain features: the fun and excitement associated with the novelty of this experience; the electronic access of the evidence survey which made it accessible at any time and at any place; the lack of pressure to perform; the chance to learn a scientific technique which had applications; the provision of food for thought; and the reinforcing of what they were learning in chemistry and biology classes. These aspects that they found valuable were all ideal conditions to promote learning in these disciplines of science.

This type of an educational activity “shows respect for students as human beings and recognizes their possession of rational abilities and potential for autonomous action” because of its acknowledgement of “students as ends in themselves rather than as means, as citizens whose fulfillment and autonomy are important goals in and of themselves” (Yarden et al., 2015, p. 66). The first act of this research activity was reading which has both a fundamental and a derived meaning: these two meanings will be explored next.

Reading as Inquiry from a Constructivist Viewpoint

Reading is an act of inquiry (Norris & Phillips, 2008). There is a fundamental sense of reading as described by Norris and Phillips when the world of the reader meets the world on paper. “The simple view of reading is defined by word recognition and information location” (p. 65). According to this simplistic view, when one knows what a text says, it is assumed that one knows what the text means (Olson, 1994). This simple view of reading does not characterize the type of reading that scientists regularly do as part of their work (Norris & Phillips, 2008). Reading is a constructive process in which the readers create meaning by forming inferential links between their background knowledge and the text that they are reading. From the constructivist viewpoint, this is one of the ways that a reader creates knowledge.

According to constructivism, which is the theoretical framework on which my study is in part based, one does not need to experiment with or manipulate the material world in order to be able to interpret data or understand the “language of science.” This stance of constructivism is well supported by my findings. Since the first part of participating in the evidence survey was to read the APL, reading was definitely treated as the first intervention into the text covered in the adapted paper. The participants performed no experiments or manipulations of the material world. The knowledge that students gained, about interpreting data and understanding the language of science used in the APL, was initially from this first act of reading.

“A core feature of science is that it is a cultural activity undertaken through the medium of language” (Osborne, 2002, p.204). Literacy is, therefore, a central requirement in science. As Norris and Phillips concluded (2001) “literacy becomes constitutive of science itself” (cited in Osborne, 2002, p. 206). “For just as there can be no house without roofs or windows, there can be no science without reading, talking and writing” (Osborne, 2002, p. 206). Consequently if we desire our students to understand and get insights into the nature and manner of scientific reasoning, we have to provide them with ample opportunities in using and exploring the language of science (Norris & Phillips, 2008). This ambitious goal can be achieved by getting students to read science, discuss what the texts mean, and argue how the ideas in what they read are supported by evidence (Osborne, 2010). The biggest barrier to students developing a deep understanding of what the texts they read mean is the vocabulary specific to these highly specialized scientific fields. In the APL that I created therefore I have attempted to minimize this barrier by providing a running list of the more complicated, scientific words used in the primary literature. These terms were underlined in the text and then their simplified meanings were provided in the column to the right. Students must also be exposed to the process of writing in and communicating using the language of science. As Yore and Treagust (2006) claimed "no effective science education program would be complete if it did not support students in acquiring the facility of oral science language and the ability to access, produce and comprehend the full range of science text and representations (p. 296).

Because of this connection between reading and science, it is "a proper goal of science education to teach students how to read papers such as these and to make it a requirement of the curriculum that they do read them" (Yarden et al., 2015, p. 66). A good way to achieve the objective of developing critical scientific literacy and the critical appraisal of scientific knowledge in undergraduate students is to introduce the reading of primary literature into their curriculum (Yarden et al.). This could be done by teaching certain portions of the curriculum using a series of relevant primary literature (Yarden et al., 2015), or having activities such as my research activity incorporated into the actual teaching of the science curriculum (Norris et al., 2009b). Scientific papers provide records of the inquiry process undertaken by scientists (Yarden et al., 2015). "Teaching students the provenance of and grounding for knowledge provides for them the first point of entry into the critical appraisal of that knowledge" (p. 66). Consequently in science, using a simpler version of the primary literature referred to as APL might make these pedagogical tasks easier. The APL that I made and employed in this study was not used for covering the curriculum, but for research purposes only.

Views on the Adapted Primary Literature (APL)

Is APL the best way to introduce a layperson to science? Why not use the reports of science in the media, because that is what the layperson encounters more often? The problem is however, that media reports very often do not sufficiently represent the epistemology of science (Norris et al., 2009a). Reading media reports provides scientific information, but these reports do not include the justification for the scientific knowledge described. One has to read the APL in order to trace the reasoning from the evidence in each context. Therefore, reading media reports cannot take the place of reading APL: it is important to read both, because each serves its own function for the novice reader. The only problem here is the time needed to cover both in class.

In my study, about three-quarters of the students stated that they enjoyed reading the APL, although a few struggled with the initial reading and found that frustrating. Students experienced difficulties as they tried to make meaning out of molecular methods used in the APL. This may be partly because the biotechnological methods are abstract and also because the students have to use prior knowledge of genetics to comprehend these methods. For students who did not have this genetics background, the initial reading of the APL was a struggle and at first some found the paper confusing. The students further found that different sections of the APL article evoked different types of challenges for them.

The argumentative genre adopted by APL is highly conducive to developing scientific reasoning. This skill in argumentation (i.e., reasoning from evidence) cannot be developed by reading textbooks since they contain mainly narrative and expository text. “By being argumentative, the APL could be interpreted as an authorization to argue, especially when accompanied by the implicit endorsement of the science teacher who is asking students to read it" (Yarden et al., 2015, p. 68). It would be fair to say that the "APL both demonstrates what good scientific thinking is like and sanctions that mode of thought" (p. 68). The argumentative genre guides my seminal research question, which is: Can Concepts of Evidence be taught or learnt; and if so, what are the factors that influence their teaching and learning? Most of the work on the usage of APL in science classrooms has been done so far in high schools. Not much has been done at the undergraduate level or in interdisciplinary science, and therefore the findings from my study would make contributions to those fields.

The APL, that the students in my study had to read initially as they participated in the evidence survey gives them a very good taste of how research is done, because it describes the objectives, the method, the equipment used and the results of the experiments done by the researchers. This gives students a very realistic glimpse of what research science is like. The APL can also "serve as a model of good scientific writing and thinking" (Yarden et al., 2015, p. 68). Besides, APL can also act as a source of inspiration because it "requires students to use their brains in ways that are novel for school" (p, 68), and it conveys an uncertainty (Varttala, 1999) which is very characteristic of the scientific endeavour. The uncertainty is brought out in what is referred to as hedging which is "the expression of tentativeness and possibility" (Hyland, 1996, p. 433) in the academic scientific discourse. Hedging depicts the tentativeness of scientific claims in such a way that it allows room for readers to expect and express possible objections to these claims (1996). In the APL used in my study, the tentativeness is clearly brought out by the authors of the primary literature acknowledging that this is preliminary work in trying to use molecular beacons to detect nucleic acid damage or alterations to just one base in the oligonucleotides exposed to uv-light. The authors of the primary literature used expressions such as “we report” and “it appears” to emphasize the fact that they realize theirs is just one interpretation of their observations and that they are open to objections to or different interpretations of their findings. In their conclusion, these authors also state “Supporting Information Available” indicating that their work is an ongoing scientific exploration and some of the supporting data not included in this paper could be provided on asking. This primary literature written by a number of scientists working together in the same laboratory, clearly illustrates that “Science is a communal rather than an individual endeavour; scientific knowledge is constructed ... by individuals in interaction with one another in ways that modify their observations, theories and hypotheses, and patterns of reasoning” (Longino, 1993, p. 11). Realizing that science has these characteristics might “make it seem less intimidating, even more human” (Yarden et al., 2015, p. 68) for first year undergraduate students.

Since I have discussed above how reading is an integral part of the inquiry process of science, I would fully support Schwab (1962) when he asserts that scientific papers are “the most authentic, unretouched specimens of inquiry that we can obtain” (p. 81). Although APL is not “unretouched,” it maintains the canonical structure of the primary literature and should consequently “provide authentic specimens of inquiry” (Yarden et al., 2015, p. 67). The APL that I created adopted the canonical structure of the primary literature very carefully. Besides, in the writing process the APL that I created was carefully checked by the chief author of the corresponding primary literature for authenticity and accuracy. In addition, there is now evidence from two studies conducted in Israel and in Canada (Baram-Tsabari & Yarden, 2005; Norris et al., 2009b) that the usage of APL promotes the scientific inquiry of students. In both studies it was shown that the students who read the APL responded more critically to the text. In my study, many of my student participants viewed the APL as a good model of scientific thinking and writing. For example, referring to the reading of the APL Student 6 was compelled to read “because it was such a pleasure and, the more I read the more I continued” (Student 6, p. 7). According to Student 9 the APL provided “a good background ... to understand the paper” (Student 9, p. 30). Student 17 confirmed “I thought it was well written, and easy to read. I learnt something as well from the Journal” (Student 17, p. 10). Student 44 “found it helpful because when you’re reading scientific papers they tend to get wordy and they don’t explain things very well and depending on how old the paper is the way it is written is really knotty or really bland so this was nice. I liked it very much ... just like how chemistry and genetics came together, two of the hardest courses out there and the stuff was so interesting” (Student 44, p. 13).

The questions that I asked especially in the interview were semi-structured and involved less enumeration of facts to remember and more provision of arguments to engage in rationally. This approach, was perceived by my students, as “far more respectful of their intelligence and thus as something to emulate” (Yarden et al., 2015, p. 68), in line with the comment by Yarden et al. Thus the APL might act just as a catalyst to inspire the participants. As Student 2 commented “May be the molecular beacons could detect damage from one total genetic mutation. ... It is huge what they (i.e. the authors of the primary literature) were thinking!” (Student 2, p. 6). When asked which part of the APL was most interesting Student 52 said that the intimacy conveyed by the APL made it fun. This student was so drawn to the APL that Student 52 could firmly say “I think it would have been fun to do the experiments myself” (Student 52, p. 38). Students are not always drawn to experiments, especially when these investigations are very involved and complicated. Therefore this is a great compliment to the APL and its objective of making sure that first-year undergraduate students taking science courses, understand, enjoy and are inspired by, this adapted form of the primary literature.

In the above section on “Reading as Inquiry from a Constructivist Viewpoint” I have described how there is a monolithic, fundamental view of reading as well as a more critical view. “The essence of the monolithic view of reading is that all reading is decoding, in the sense of figuring out what the words are” (Yarden et al., 2015). The simple, fundamental view of reading views all scientific texts as having a predominant feature – the presence of scientific vocabulary – words that refer to particular scientific concepts, processes or methods (Norris et al., 2008). One cannot deny that there is a lot of scientific vocabulary in science texts. “However that vocabulary does not have to be the dominant feature of those texts” (Yarden et al., 2015, p. 112). It is for this reason that I have attempted to provide the meaning of all such terms used in the APL that I have adapted. Then the APL can be used creatively to draw attention to many other features of science other than its vocabulary. It is this simple view of reading in combination with the preponderance of scientific terms in school science texts that lead to the view that science is difficult to read.

If the critical reading of science has to replace this simplistic view of reading, then science education has to confront and remedy two unflattering truths (Yarden et al., 2015). First, many high school graduates cannot read science very well if reading is meant to signify the critical reading of science texts. Second, students are not taught to read science in high school or at the first year undergraduate level. Very little instruction, if any, exists in reading scientific texts in high school or at the undergraduate level (2015). Instead of focussing excessively on scientific vocabulary, the alternative may be to place emphasis on scientific language that is easy to recall and that is useful for those who do not want to make a career out of science but who still want to remain scientifically literate citizens. One of the productive ways to proceed with teaching reading in science at the high school and at the undergraduate level may be to focus on three aspects of scientific texts: text structure; text epistemology and meta-scientific language (2015). As illustrated in detail below APL in general is poised to amply provide all three aspects in good measure.

Text structure. Three types of text structure have been identified: organizational, goal-directed and argumentative. All three are evident in the APL that I created. The organizational structure is what provides the reader with an idea of the location in the paper. For example, maintaining the canonical structure of the primary literature, my APL is also divided into sections such as introduction section, methods, results and discussion. Students could be taught the "speech acts" (Yarden et al., 2015, p. 119) to expect in each section. For example, reporting was prevalent in the methods section of my APL; while findings were explained in the results section; and speculating entailed a good part of the discussion section. "Teaching students how to identify these structural signals in text and how to adapt their stances accordingly is fundamental to learning to read scientific text" (p. 119).

The goal-directed structure is more involved than the organizational structure. Here students have to be taught to identify both the overall goals of the whole study as well as the local objectives of one section of the APL such as the methods section (Yarden et al., 2015). In my study the students' understanding of the overall goals was tested in my interviews when I asked them to describe the overall aim of the various experiments conducted in the paper as a whole. One of the local objectives of the methods section was to describe the intricate details of how the concentration of the oligonucleotides was determined by calculating their molar extinction coefficient and their absorbance. If students miss the goal-directed structure of the text, they will miss the whole point of the writing (2015).

The argumentative structure forges links between the primary evidence that is reported in the paper and the major conclusions reached therein (Yarden et al., 2015). For example, the results and discussion section claimed from the empirically–obtained data plotted as Figures 6 and 7 in my APL, that the fluorescence always decayed three times faster than the absorption and the fluorescence and absorption decay rates were greater for the rU_{17} than for the dT_{17}. Each of these interpretations poses challenges which could be addressed if reading instruction was implemented at the high school and first-year undergraduate level by introducing the reading of APL to supplement what is taught in the curriculum.

Text epistemology. All scientific papers provide data, reasons and evidence for conclusion which constitute their argumentative structure. Papers in science also "express a degree of tentativeness in their findings" (Yarden et al., 2015, p. 49). These two characteristics specify distinctly the epistemology on which these papers are based. This epistemology is termed "fallible rationality" (p. 49). It is not possible to understand a scientific text without understanding the epistemology expressed in it either explicitly or implicitly. Fallibility was expressed outwardly in my APL by using expressions such as "may" or "might," verbs such as "appears" or "suggests" and adverbs such as "perhaps" or "possibly." Therefore, it is important to teach students the variety of ways in which a text can be used to express rationality and fallibility. Interpreting text epistemology is not achieved easily or quickly. Therefore although the ability to interpret the epistemology of the text may start at high school, it will continue to develop as students move into the first-year undergraduate level.

Meta-scientific language. Those who speak and write science are able to refer to scientific practices in a theoretical and critical way by using meta-scientific language. By using meta-scientific language, historians, philosophers and sociologists of science are able to take a stance as observers of science. In other words, they take the position of observing science from the outside-in although they are insiders to their own research. The meta-scientific language used in my APL includes terms such as “observe,” “data,” “results,” “effect,” “predict,” and “explain.” The three initial terms “observe,” “data,” and “results,” have to do with evidence in one way or another, and draws attention to the rational epistemologies of the authors. Every one of the above-mentioned, meta-scientific language terms “implies a relationship and an evaluation based on that relationship” (Yarden et al., 2015, p. 53). For example, when we refer to a phenomenon as “observed” it is equivalent to evaluating it as suitable for furnishing evidence. Students need to learn to notice the usage of meta-scientific language when they read scientific texts because if they fail to do so, they will not grasp the full meaning of those texts. All three aspects of scientific texts can be taught successfully with direct instruction, modelling on the part of the science teacher and lots of practice.

The Concepts of Evidence identified within the APL highlight the fundamental details of research science and accentuate the aspects that students are aware of from their participation in science labs. This clearly shows them what they learn and do in the laboratory as normative science, has applications in research. It also demarcates and draws attention to the differences between research science and normative science. APL serves as an excellent vehicle, therefore, to understand this distinction. This is the first study where APL has been used as a medium to measure the understanding that students have of Concepts of Evidence.

The Teaching Intervention and Pedagogy

The students were greatly appreciative of the fact that the teaching intervention was made available to them on-line because this meant they could access it any time and watch the video as often as they needed to, in order to fully understand the material taught. Besides, in the present age, the audio and visual components of the video may be preferred by first-year undergraduate students over the live teaching by an instructor. In addition the conversational style of the teaching intervention was appealing. The presenter faced the whiteboard, as he talked and either wrote or drew on it at the same time. The dual sensory stimulation (i.e. by sight and sound) reinforces the learning: the greater the sensory stimulation the deeper is the learning (Zull, 2004). Personal anecdotes that the chemistry professor interspersed into the video made the teaching intervention more intimate and meaningful: this was only possible because the professor was the chief investigator in this research project. The visual and anecdotal explanations were lauded as pedagogically sound in successfully linking empirical evidence to the underlying scientific theories.

Overall, most students were highly appreciative of the teaching intervention and found it very helpful in enabling them to understand the APL as well as the Concepts of Evidence nested in them. An important factor in favour of the teaching intervention is that the video lasted for only 32 minutes which is what Wilson et al., (2007) note as the attention span of most first-year undergraduate students. However, according to these authors the concept of a single attention span is not well-founded. Furthermore, recent research has shown that the attention span decreases as the lecture progresses (Bunce et al., 2010). Therefore, if the students found that their attention span was decreasing, they could stop the video at any time and come back to it when they were more refreshed and better able to concentrate.

Students commented enthusiastically about the in-built advantage of this on-demand access to the lecture material. The teaching intervention could be tailored so that the length and pace of the lecture was better suited to a student’s own engagement with the material in a way that fostered understanding. This feature could also enhance note-taking. Research has shown that first-year undergraduates record only 10% of the important points in a typical lecture (Locke, 1977). Therefore the fact that the video could be stopped at any point in the lecture meant that students could note down the main points better and this could enhance the learning and understanding of both the content and the procedural knowledge. These in-built considerations highlight how the teaching intervention is a very appropriate pedagogical tool.

The positive comments from students about the efficacy of the teaching intervention were very encouraging. Although there are 21 categories of Concepts of Evidence according to Gott's classification, the chemistry professor dealt with only four of them in the video because I had listed them as the major Concepts of Evidence implicitly covered in the APL. He described these Concepts of Evidence as the underlying formalism (i.e., logical structure of a scientific argument) of how experiments are designed to display certain characteristics. The four Concepts of Evidence he covered were: "Instruments: Calibration and Error"; "Sampling a Datum"; "Variables: Fair Test"; and "Patterns and Relationships in Data." The chemistry professor referred to the diagrams and graphs already drawn on the whiteboard to describe when and how to use these Concepts of Evidence. In this way the teaching intervention stressed the fact that the "making sense of science is a dialectical process involving both content and process. The two can never be meaningfully separated" (Tobin, 1993 p. 9). It was evident from the teaching intervention that the procedural knowledge (i.e. Concepts of Evidence or process) was embedded in the content knowledge. To understand the Concepts of Evidence, therefore, it was essential to have a fundamental understanding of the content knowledge. Conversely, a deep understanding of the Concepts of Evidence enhanced the substantive knowledge as well.

The role of the Concepts of Evidence is to stipulate the factors controlling the collection, measurement and handling of data. Having the pre-test made it possible for the students to know and to measure their initial substantive knowledge. This is consistent with the constructivist pedagogy which takes into account what students know and scaffolds the teaching to build on that initial knowledge. The scaffolding in my study is provided by the teaching intervention which had tremendous efficacy in furthering the students' understanding of both the substantive knowledge of the APL and the procedural knowledge (i.e. the Concepts of Evidence embedded in the APL).

Roth and Roychoudhury (1993) defined scaffolding in the following way:

> When an adult, an expert, or a more capable person facilitates the performance of student so that achievement is at a higher level than was achieved individually, the process is known as "scaffolding." ... The concept of "cognitive apprenticeship" has emerged to describe the learning through expert-guided experience on cognitive and metacognitive skills.
>
> (p. 133)

The teaching intervention has a scaffolding role because it is presented by an expert in the field of chemistry, photobiology and photochemistry. The pre-test of the evidence survey revealed the understanding that students were able to achieve on their own with the reading of the APL. Since the performance of 86% of students was better on the post-test it is evident that the teaching intervention facilitated the performance of these students, and the presenter was able to guide these students to achieve a higher level of understanding through "scaffolding." Scaffolding, promoted by constructivism, is a style of instruction that provides students with the intellectual support to function at the cutting edge of their individual development. Students need to be supported in their learning. The two categories of students who need the most support are those who are at risk for failure, as well as strong students who are working to master difficult materials.

It is only skilled experts who are able to offer this type of "cognitive apprenticeship" or support. Such pedagogical skill requires not only content knowledge but the skill to communicate and understand the needs of the students with whom they interact. These two together are referred to in the literature as pedagogical content knowledge (PCK). The presenter of the teaching intervention in my study is a 3M National Teaching Award (Canada's most prestigious recognition of excellence in educational leadership and teaching at the university and college level) holder. By utilizing the constructivist pedagogy in the teaching of the Concepts of Evidence, the presenter was able to "build knowledge structures which are commensurate with the knowledge of the discipline" (Tobin & Tippins, 1993, p. 7). In the teaching intervention, therefore, this expert was able "to provide students with a scaffold to build knowledge in directions that would not be possible without the influence of a teacher" (p. 10).

A very important pedagogical factor to consider is that the teaching intervention was not test-score-driven or curriculum-based. Students had to take ownership for their own learning but there was no pressure to perform. That is why I decided not to share their pre-test and post-test scores with the students. In this way the formative assessment was kept low-key and not emphasized. Learning was designed to be for fun, so that the goals of deriving meaning and understanding which are implicit to constructivism were met. Very often society places a virtually exclusive emphasis on test scores. When this emphasis is transmitted to the students it becomes more and more difficult for teachers to pay less attention to the assessment and watch more carefully the actual learning process. "The success of constructivism has come from reminding teachers that children do have ideas and theories and that their thinking is the foundation on which new meanings must be formulated" (Osborne, 1996a, p. 65).

Teachers have to build on the existing neuronal networks of their students. In my study this initial neuronal network is revealed through the pre-tests. Knowledge resides in the brain as neuronal networks (Zull, 2004). Repeated usage of the neuronal networks, lead to the branching of neurons and the formation of interconnections between neurons which increase the density of that region of the brain (2004). This is what happens in the brain when students build on their prior knowledge. “These networks are the physical equivalent of knowledge, and the change in the connections that make up the networks is learning” (p. 70). With the modern advances in research in education and cognitive science, we now have deeper insights into the changes that happen in the brain when we learn and remember.

The Semi-Structured Interviews and Scaffolding

As I interviewed the students who participated in my study and marked their pre-tests and post-tests, I was keenly aware that not all students had the same level of understanding of the substantive or the procedural knowledge. Although the pre-test and post-test of the evidence survey shed some light on the understanding of substantive knowledge and procedural knowledge held by individual students, the interviews were a much better gauge of the exact level of understanding and of any alternate conceptions students may be harbouring. This is because the interviews were a one-on-one interaction and since they were semi-structured in nature it was possible to ask good, impromptu questions which could shed more light on what exactly each student had understood.

The psychologist, Vygotsky (1978), stated that cognitive development stems from social interactions especially those providing opportunities for guided learning. During the interview in my study each student was given the opportunity to interact with “a more knowledgeable other,” which was me. Each semi-structured interview was adapted so that there was an ongoing diagnosis of student learning. I guided their learning by asking probing questions and filling in the gaps and correcting the errors or misconceptions in their knowledge. Such ongoing support offered to a learner by an expert was termed “scaffolding” by Wood, Bruner and Ross (1976). These scholars defined scaffolding as an “adult controlling those elements of the task that are essentially beyond the learner’s capacity, thus permitting him to concentrate upon and complete only those elements that are within his range of competence” (p. 89). The notion of scaffolding was formulated “to explore the nature of the support that an adult provides in helping a child to learn how to perform a task that, alone, the child could not master” (Wood & Wood, 1996, p. 5).

One of the most crucial aspects of this notion of scaffolding is the role of the adult or the expert (Puntambekar, 2009). Here the expert not only has a thorough knowledge of the content of instruction, but has the skills, strategies and processes needed for teaching the topic (2009). Since I spent time reviewing the substantive knowledge entailed in the APL and the Concepts of Evidence as laid out by Gott et al. (2003), I could be classified as an expert in the content knowledge covered in this research activity.

Although the role of the expert is critical, instruction which is scaffolded effectively has the following key components and goals that were observed in my research activity:

(1) Common goal or intersubjectivity: There is a combined ownership of the task at hand (Puntambekar, 2009). In my study each student and I had the common goal of answering the interview questions, and completing the research activity at hand. Although each student and I began the interview with a different understanding we usually arrived at a shared understanding by the end of the interview session. In the literature, a similar arrival at a shared understanding is illustrated in the work of Berk and Winsler (1995).

(2) Ongoing diagnosis and adaptive support: The most important feature of scaffolding is that the expert is constantly evaluating the student’s progress, and providing support that is appropriate (Hogan & Tudge, 1999). The semi-structured nature of my interviews

enabled me to ask timely questions to fathom the students' understanding, to diagnose any misconceptions they may be harboring and to support them in the mastery of the content and procedural knowledge covered by the questions.

(3) Dialogues and interactions: "A critical factor in the ongoing diagnosis and calibrated support is the dialogic nature of scaffolding interactions" (Puntambekar, 2009, p. 2). During the interviews, the students in my study participated actively, contributing to the direction of the dialogic interaction, and were not passive recipients in any sense. The questions that I had prepared for the student interviews were very useful in maintaining the dialogue and not straying from the task at hand.

(4) Fading and transfer of responsibility: The last feature of scaffolding relevant to my study is that the support provided by the expert is weaned off as the interaction progresses so that the learners are in control and take over the onus for their learning (Puntambekar, 2009). This was very evident in my interviews because I initially explained whatever needed clarification but gradually the students were able to answer the questions right away without any further clarification or explanation on my part.

I fully agree with Phillips (1995), "... it is the cognitive effort of the individual that results in the construction of knowledge" (p. 8). The degree to which students are prepared to "wrestle" with the new content knowledge or the newly-introduced Concepts of Evidence decides their level of understanding. There is a cognitive struggle involved to fully understand the new material and fit it into the knowledge base one already possesses. This is in full agreement with constructivism which does not endorse the "spectator theory" of learning where one learns by watching or being a spectator (Phillips, 1995). According to constructivism the learner has to be an active participator or actor in the learning process and only then does deep learning occur.

The teacher and students together co-create the new knowledge, and each has an active role to play in this co-construction.

The theory of cognitive development proposed by Lev Vygotsky (1978) focussed also on the role of culture in the development of higher mental functions. One important concept that arose from this theory is the zone of proximal development, commonly referred to as ZPD. According to Vygotsky the ZPD is defined as "the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance, or in collaboration with more capable peers" (Vygotsky, 1978, p. 86). The ZPD signifies the range of tasks that a learner can perform with the help and guidance of others but cannot yet perform independently. This zone may also be thought of as “the region in which this transfer of ability from the shared environment to the individual occurs” (Berk & Winsler, 1995, p. 24). Within this zone of proximal development, there are two levels: first the actual development level, and second the level of potential development. The actual development level is the upper limit of tasks that one can perform independently. In my research activity the pre-test of the evidence survey measures the initial, maximal level of understanding of the APL and the Concepts of Evidence attained by the students independently: this is a measure of the actual development level. The second level known as the level of potential development is the upper limit of tasks that one can perform with the help of a more competent individual. The post-test of my evidence survey provides a measure of each individual student’s level of potential development, attained after participating in the teaching intervention offered by a more competent individual (i.e. the instructor who delivered the teaching intervention).

When I interviewed the students, I had already marked their pre-tests and their post-tests.

Therefore I directed a part of each interview to address the flaws each student made in the pre-test and the post-test. According to Vygotsky (as cited in Coffey, 2009), the zone of proximal development is the area “where the most sensitive instruction or guidance should occur” (p. 1). Teachers, parents, and mentors attuned to a learner can recognize where this student is within the zone of proximal development by asking questions (Berger, 2009). I had the opportunity to promote the growth in the understanding of individual students within their zones of proximal development while I interviewed them.

Dynamic Interrelationship between Teaching Intervention and APL

The constructivist approach emphasizes the fact that if learning is to occur, the learner must reconstruct knowledge during the learning process. In this reconstruction, learners are able to appropriate the new knowledge as their own. In my study, the extent of this reconstruction depends on the extent of the interrelationship between the APL and the teaching intervention. If this inter-relationship is strong and dynamic, it will promote the individual reconstruction of pre-existing knowledge.

The erroneous conceptions that students have are based on their mental models of how the world works (Chi, 2000a). Sometimes students’ prior experiences and the ways they understand them, make it challenging for them to learn the content and techniques their instructors want to teach them. Students bring their own mental models of Concepts of Evidence as well when they read the APL and answer the pre-test. Then, they attempt to slot the new information from the APL into their existing mental models which are flawed. This just does not work and may not always be pleasant but may accentuate a struggle as the old mental models try to slot in the new information. As a consequence of this struggle, students will be required to reconstruct their initial mental models and form new ones: their mental models may have to undergo a conceptual change.

Conceptual Change

The ground-breaking work in the field of conceptual change was carried out by Posner, Strike, Hewson, and Gertzog (1982), when they tried to develop a model for this phenomenon, and revised their model a decade later (Strike & Posner, 1992). Posner et al. (1982) support the constructivist view that, before coming to the classroom and receiving any pedagogical intervention, students do not come in as blank slates, but possess their own ideas about scientific phenomena. These initial ideas that students have may or may not be scientific, and have a significant influence on the way students learn new scientific concepts through the teaching intervention (Çetin, Kaya, & Geban, 2009). Learning is achieved through the interaction between new knowledge and one's existing, preliminary conceptions (Hewson, 1992), and the final outcome of learning is determined by the nature of this interrelationship.

In several topics on the school curriculum, students possess these pre-instructional conceptions which hamper the way they learn scientific concepts taught in school (Chinn & Brewer, 1993; Driver, Squires, Rushworth, & Wood-Robinson, 1994). Often these pre-instructional conceptions, alternate conceptions, or erroneous conceptions as they are referred to sometimes, are very resistant to change and impede conceptual change in science. This has been found to be true with children, adolescents and adults on a variety of topics in science, and in a wide range of cultures and ethnicities (Chinn & Brewer, 1993, 2000; Driver et al., 1994; Duit & Treagust, 1998).

Erroneous Conceptions

The conceptions that people have which are "different from the scientifically correct ones" (Çetin et al., 2009, p. 130) are variously termed misconceptions, alternative frameworks, children's science, preconceptions, and alternative conceptions (Çetin et al., 2009; Driver & Easley, 1978; Nakhleh, 1992; Nicoll, 2001). According to Nakhleh (1992), misconceptions will interfere with students' subsequent learning. It is therefore vital to have concept-change-oriented instruction geared towards addressing some of these erroneous conceptions, so that the thinking of students who hold these unscientific views can undergo conceptual change, and eventually lead to cognitive development. One of the chief aims of science education, consequently, must be to help students attain the correct scientific understanding of important science concepts. Table 2 shows some of the erroneous conceptions held by the participants about the fair test (CoE #12), which became evident during this study. When these preconceptions and misconceptions about the fair test are known, the teaching intervention could be tailored to address them. This would give students who possess these erroneous conceptions a chance to face and rectify them because these alternate conceptions get in the way of deep understanding.

Table 2

Erroneous Conceptions of Fair Test (CoE #12) Held by Participants in this Study
Doing a control experiment makes a fair test.
All confounding variables were accounted for.
A fair test is one using a different substance and a different wavelength.
When there is a controlled variable it is a fair test.
A fair test is when data are manipulated to the greatest advantage.
When all the conditions in an experiment are stated then it is a fair test.
Time of irradiation is the responding (dependent) variable.
A closed system yields a fair test.

The manner in which students organize information matters, and helps them develop expertise. As new information is assimilated, new nodes will be formed, nodes (both new and pre-existing) will be interconnected, resulting in webs of interconnected knowledge. Initially in novices, the connections between the nodes will be sparse and superficial (Chi, Feltovich, &

Glaser, 1981). In experts, however, the connections increase in number, and large webs of inter-connected knowledge are produced. These inter-connections result in meaningful knowledge. It is not easy for students to change from novices to experts, and they need to be supported in this effort (Chi et al., 1981). In my study, this support is provided by the teaching intervention as well as the interviews.

A reorganization of the new knowledge in the APL, therefore, occurs as a result of watching and listening to the teaching intervention. If there is a strong and dynamic relationship between the APL and the teaching intervention, the latter will not only explain, but take on the role of organizing the information and knowledge in the APL. When the teaching intervention is able to give the knowledge in the APL an overall structure (i.e., sequential or hierarchical), it will aid in the understanding, recall, and retention of this information (Chi, 2005). For example, the explanation of sampling offered by the chemistry professor in the teaching intervention used in my study, took on the form of a sequential flow of events, while the Concepts of Evidence which are arranged into categories and sub-categories, were described as having a hierarchical structure. New nodes will be created and novel interconnections between new and pre-existing nodes will occur readily in the brains of students as the structure of the new information unfolds. This will lead to more inter-connected nodes and more organized webs, producing more meaningful knowledge in the minds of students. Consequently the information in the APL will be slotted into the appropriate positions in the new mental models formed by the students, promoting the understanding, the learning and the recall of this knowledge. This process in its entirety brings about a conceptual change.

The Processes and Challenges of Handling Erroneous Conceptions. Although some of the students in this study had to struggle to understand the APL, I was overwhelmed, as I interviewed them, to realize that some students still harboured many inaccurate conceptions. The naive knowledge that students have are termed “preconceptions” while those misunderstandings that persist strongly even after instruction are called “misconceptions.” As I have a deeply invested interest in teaching and learning, I was determined to explore how to handle misconceptions.

Misconceptions can be envisaged as concepts categorized into an ontologically inappropriate category. They are often naive qualitative explanations that are flawed. For example, the belief that “Doing a control experiment makes a fair test” is an explanation that is flawed. Having all the variables controlled in an experiment does not make it a fair test. The distinguishing characteristic of a fair test is that one independent variable is altered at one time and all other variables are controlled. Conceptual change, therefore, is merely the process of reassigning or “shifting” a miscategorized concept from one ontological category to another. In the example above, the fair test is shifted from the “controlled experiment” category to the “One independent variable varied at one time” ontological category. A good fundamental understanding of procedural knowledge is vital in this process of “re-slotting” into the appropriate category. This is where teaching about Concepts of Evidence and developing a fundamental understanding about these concepts is important. The process of repairing misconceptions is termed *conceptual change*, while the means of repairing preconceptions is called *conceptual reorganization.* In spite of the fact that conceptual change has been discussed over several decades, the literature paints a blurred picture of what exactly constitutes conceptual change and why it is difficult. This is what I want to address next.

As science educators, if we are aware of common student misconceptions in the topic we are teaching, we can correct those misunderstandings. However finding the misconceptions harboured by a particular student will require individual attention such as that provided in my study by the student interviews. There are no “standard misconceptions” for each topic in science. Teaching interventions could be tailored to address and eliminate the misconceptions which are identified. The results of the analysis of the evidence survey clearly demonstrate that pedagogy does have a significant role in improving the understanding that students have of Concepts of Evidence in science. Very often, however, science educators teach a topic without a clear idea of students’ misconceptions, and what frames these misconceptions fit into. If students with misconceptions are given a teaching intervention addressing these alternate conceptions, they will be shown the error in their understanding. They will have to then face the error and correct their own misunderstanding. This may result in a conceptual change. I cannot help wondering what is the mechanism invoked in these conceptual changes. It is also intriguing to explore how and why conceptual changes are hindered.

According to Posner et al.’s theory (1982), four conditions have to be met for conceptual change to occur in an individual’s understanding: (1) There must be dissatisfaction with existing conceptions; (2) The new conception offered must be intelligible; (3) The new conception must be initially plausible; and (4) The new conception must offer the possibility of a fruitful research program. I have illustrated with detailed examples, in Appendix J, how the four conditions required for a conceptual change to occur have been met in some of the interviews with my student participants.

A conceptual change only occurs when a person engages with the text of the APL and, on questioning through face to face interviews, displays a change and an improvement in the understanding of the Concepts of Evidence. Authenticity is not achieved by merely “engaging with ... materials that others have assessed to be ‘authentic’” (Osborne, 2009, p. 400).

Authenticity has to be earned or experienced and "results from a commitment by the individual to seek understanding and purpose in any activity" (p. 400). The fact that a conceptual change occurred is evidence of authenticity. However, Osborne cautioned that authenticity cannot simply be attributed to a particular activity.

In this study, students who have displayed a conceptual change have a heightened level of comprehension. This may be because they have experienced dissatisfaction with their existing conceptions creating a disequilibrium which they have had to face (Read, 2004). Consequently they developed a willingness to accommodate a new concept which seemed intelligible and plausible. Osborne (2009) maintained that just introducing APL materials into the science classroom does not guarantee that authenticity will be the experience of each student: each student has to share this view. Just because the text is similar to what scientists use does not make the experience authentic, but all students are required to make the effort to increase their understanding. During the interviews in my study, some of my student participants obviously displayed conceptual change when the scaffolding was removed.

The process of achieving a conceptual change is not inherently difficult, but can prove to be challenging when students are not aware of their misconceptions and they lack the alternative ontologically distinct categories to which they should shift their misconceptions (Chi & Roscoe, 2002). It is the second obstacle that poses the greatest problem in repairing certain misconceptions. Some processes in science are causal and sequential such as the effect of heat on the molecular beacons. In other processes such as the effect that time of irradiation with uv-light had on a mixture of molecular beacons and DNA/RNA, students had to consider the collective interactions of all the molecules. In other words, to predict the outcome, students had to consider the interaction between molecular beacons and DNA or RNA. Such processes are, therefore, referred to as non-sequential and *emergent.* In such emergent processes, students have to focus on the sum of all the collective interactions of the individual molecules across time. In order to interpret and learn such non-sequential, emergent processes, students need to use a complex "emergent process schema," in both formal and informal contexts (Chi, Roscoe, Slotta, Roy & Chase, 2012). Such an "emergent schema" might be absent in most students and this may account for why they have robust misconceptions. It is important to teach students what an "emergent schema" is and how to handle such schema. This will equip students better to learn and understand science processes which are emergent in nature, such as diffusion and evolution.

Attempts to account for the mechanism of conceptual change have adopted two main theoretical approaches: the coherence view and the fragmentation view (Sengupta, Krinks & Clark, 2015). Lately, Sengupta et al. (2015), have tried to develop a framework for analyzing conceptual change qualitatively (Sengupta et al., 2015). The literature on conceptual change, however, shows that there are some common results that most researchers in this field would agree upon. First, "the task of conceptual change requires effort on the part of the learner" (Read, 2004, p. 17). Second, pre-existing knowledge plays a crucial role in attaining conceptual change. Third, conceptual change is a gradual process. Beyond this point opinions diverge and there are mainly four schools of thought each bearing its own implications for instruction (2004): they are Vosniadou's view; Chi's view; diSessa's view; and the Sociocultural view.

According to Vosniadou conceptual change takes place in the mind. Prior knowledge can act as both an obstacle which needs to be overcome and as a vehicle through which change is brought about. This view was adopted in my teaching intervention, where the instruction tried to address the incorrect ontological and epistemological assumptions that students held in the context of the APL (p. 13). Chi's view asserts that conceptual change is the re-slotting of concepts into their correct ontological categories which takes place in the mind. Prior knowledge is an obstacle here because it includes concepts which have been incorrectly categorised. In my study, Concepts of Evidence such as variables were categorised incorrectly in some students' minds. For example, in Figure 6 of the APL, the fluorescence intensity may be incorrectly categorised by some students as the independent variable. Then instruction, according to Chi, must promote the understanding of the correct categorisation by teaching that the variable which is observed, such as the fluorescence intensity, is the dependent variable. In keeping with the view held by diSessa, conceptual change involved the organization of pieces of knowledge in the mind into a coherent whole picture. This was often achieved in my study during the interview phase, when questions were asked, and instruction tailored "to promote the development of mental coherence" (p. 13). In accordance with the Sociocultural view, conceptual change is not only the cultivation of cultural tools but learning when to use them and in what manner. "Conceptual change occurs in society and is socially negotiated within a community of practice" (p. 18). According to this view, instruction in my study for example, should be geared towards teaching students to distinguish between the diverse contexts in which the concepts of evidence can be used, so that learners can decide which of these concepts are useful in which situations.

Learning does not need to be through a conceptual change. If only preconceptions are involved, students add new concepts through further experience to phenomena that already exist. This type of conceptual reorganization, where the student's ideas grow and change, is referred to as *assimilation.* For example, in Table 2 (page 261) the sentence, "When there is a controlled variable it is a fair test" is correct but that is not the distinguishing character of a fair test. It is when there is only ONE independent variable that you set up a fair test. Therefore, in order to be accurate one has to assimilate and add on the knowledge that fair tests can only be set up if one independent variable is altered at one time. In other cases, however, the concepts that students currently hold are inadequate to allow them to fully understand the new phenomenon accurately. Consequently, there will be a reorganization of existing concepts, leading to the rejection of some notions and the replacement of others. This is a more radical form of conceptual change referred to as *accommodation*. For example, in Table 2 (page 261), the statement “A closed system yields a fair test” requires shifting the concept of fair test from the wrong, ontological category of “closed system” to the correct ontological category of “open system.” Then the open system in this case must be specified as one in which only one independent variable is varied at one time: only then is the process of accommodation accomplished.

How can conceptual change be impeded?

When dealing with anomalous data, it has been shown that both children and adults have resisted changing their scientific conceptions (Chinn & Malhotra, 2002). Further research was undertaken to explore the actual cognitive processes underpinning the resistance (2002). It was established that when evaluating anomalous results, four cognitive processes could be identified: observation, interpretation, generalization, and retention (Brewer & Lambert, 1993). Although theoretically conceptual change could be hindered during any one of these processes, Chinn and Malhotra (2002) found that “conceptual change was blocked most strongly” at the observation stage (p. 327).

Observation has been shown to be a crucial process in the evaluation of data (Brewer & Lambert, 1993). In the student interviews, one of the questions I asked was: “Is there sampling in the experimental set-up of Figure 6 in the APL?” If the student answered yes, I asked why they thought so. If the students are to answer the question correctly and if conceptual changes are to occur in their understanding, they must observe the following:

1. The points in the experiment are very close to each other, representing a very small span of time between them. Therefore, these points could not be separate experiments.
2. What are shown as open or closed circles must be aliquots taken from the vial holding the original hybridization mixture.

If students fail to observe the above two points, they will have no reason to conclude that there is sampling, and it is very easy to fail to observe them.

Many philosophers and sociologists of science have argued that "observations are theory laden" (Brewer & Lambert, 1993, p. 328). This implies that the observations one makes are almost completely determined by the prior theories that one holds. If this is true, then conceptual change will be impeded at the observation stage, because the time interval between the circles in this experiment is so small that the students will have to be observant to see that the experiment is depicting small samples of the original hybridization mixture. If they fail to make this observation, these students will fall back on their own initial theory, however false it may be (Champagne, Gunstone, & Klopfer, 1985). Of the four cognitive processes listed above, the one at which conceptual change is impeded most strongly was observation. The reason why conceptual change did not occur was mainly because students who have incorrect expectations are unable to make the correct observations. The other three processes – interpretation, generalization, and retention – "together produced less impedance to conceptual change than observation did alone" (p. 340).

This study accentuates the opinion that "when sensory stimuli are ambiguous, prior expectations exert a powerful influence on what is perceived" (Chinn & Malhotra, 2002, p. 328; Brewer & Lambert, 1993). Therefore, when attempting to bring about a conceptual change in their students' understanding, it is important that science educators choose experiments or demonstrations where the sensory stimuli are dependable and not ambiguous. I will illustrate this further with an example.

Earlier in this chapter, in Table 2, one of the misconceptions under fair test is “all confounding variables were accounted for.” This does not really define a fair test because the student has not specified what the “confounding variables” are. The misconception is that it is not enough to just *account* for the variables: the variables have to be *controlled*. In a fair test it is imperative that the dependent and independent variables are clearly laid down. Further, it is vital to ensure that in an experiment only one independent variable is changed at one time. If this is not done, then the experiment is not valid. When you do not do a fair test, there is no sensory stimulation to alert you that you are wrong: the sensory stimuli are ambiguous and non-committal. Therefore, students fall back on their initial, pre-instructional conceptions about what a fair test is. When these pre-instructional understandings are wrong, they invalidate the experiments based on them. Fortunately, teaching interventions can be tailored to address and eliminate these misconceptions. Therefore, pedagogy has a definite and vital role in improving the understanding of students about Concepts of Evidence.

Using descriptive drawings to promote conceptual change.

Within science education research, it is well established that some students have great difficulty when learning scientific concepts (Edens & Potter, 2003; Posner, Strike, Hewson, & Gertzog, 1982). Research has shown that students often harbour incorrect or fragmentary pre-instructional conceptions about the world, which are markedly different from scientific conceptions (Henriques, 2002; Osborne & Wittrock, 1983). The conceptual change approach to science education attempts to help learners change their “existing inaccurate or incomplete mental model that conflicts with the scientific model, to a more adequate one" (Edens & Potter, 2003, p. 135).

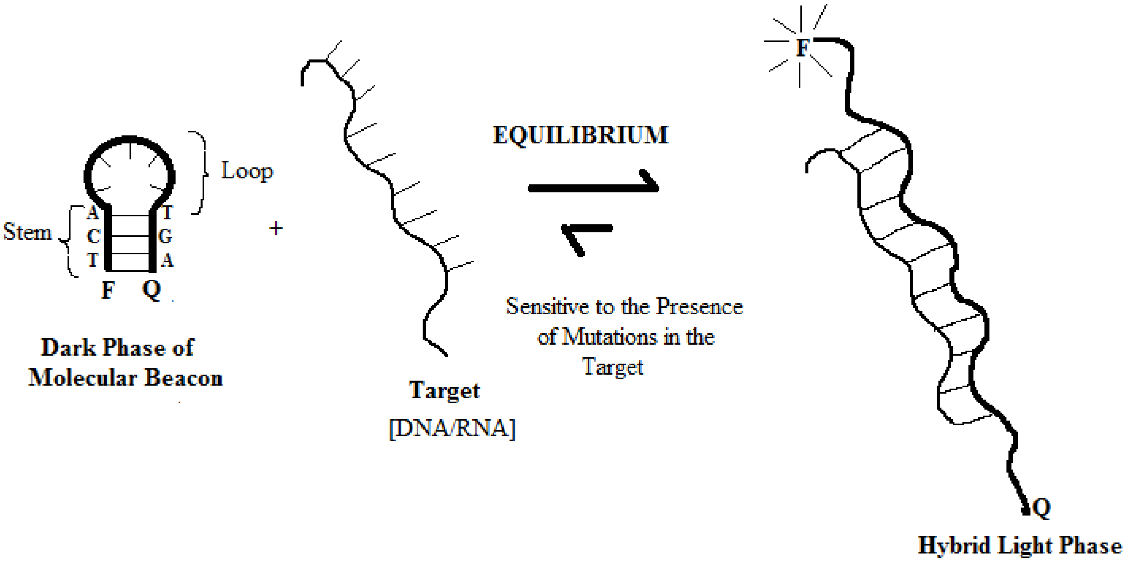
One of the methods employed to promote conceptual change is to use descriptive drawings to complement the teaching intervention geared towards correcting misconceptions. There is a strong theoretical basis for using "pictorial representation as a vehicle for promoting conceptual change" (Edens & Potter, 2003, p. 136). The dual-coding theory by Paivio (1990), states that cognitively there are two separate subsystems which handle two different classes of phenomena. One subsystem specialized for dealing with language is referred to as the verbal system (1990). The second subsystem is responsible for the generation of mental images and is called the imagery system (1990). Functionally these two subsystems are interconnected "so that activity in one system can initiate activity in the other" (p. 54). Therefore, when the verbal system and the imagery system are stimulated or coded simultaneously, a greater understanding is observed. This dual coding was attained in the teaching intervention, as diagrams were drawn to illustrate what was being taught verbally, and this promoted deeper understanding.

The teaching intervention in my study, for example, was greatly appreciated by my student participants. A significant reason for its success was the fact that the presenter (the chemistry professor) used diagrams to supplement his explanations. One instance when this was done effectively was when he described the use of the irradiator and the process of sampling. He drew a diagram of the irradiator illustrating how the hybridization mixture was placed and irradiated, how the sampling was done, and how the aliquots withdrawn from the hybridization mixture were prepared for fluorometric and spectrophotometric analysis. This type of double stimulation of the auditory and visual senses was very effective for enhancing the understanding of the substantive knowledge needed and the Concepts of Evidence embedded in the APL.

Another instance where diagrammatic representation was used effectively in the teaching intervention was in describing the structure of molecular beacons along with how and why they fluoresce. Figure 7 illustrates the molecular beacons in the free state having a hair-pin structure with a stem and loop. In the dark phase, the fluorophore and quencher are in close proximity, as shown in Figure 7, and the quencher effectively quenches the fluorescence so that the molecular beacons look dark. If a target (i.e., a segment of DNA or RNA) is added to the molecular beacons in the dark phase, the bases in the target will base pair with the bases in the loop of the molecular beacons, forming a hybrid. This will open out the molecular beacons, as shown in Figure 7, and separate out the fluorophore and quencher, positioning them at opposite ends of the hybrid. The fluorophore, which is no longer close to the quencher, will fluoresce instantly like a switch turning on a light bulb. This is referred to as the hybrid light phase. The dark and light phases of the molecular beacons are in equilibrium. All this explanation about the structure of the

Figure 7

The Dark Phase and Hybrid Light Phase of Molecular Beacons



molecular beacons in the dark phase and the base-pairing when combining with a target to form the hybrid light phase could be conveyed most lucidly by incorporating diagrams such as Figure 7. As proclaimed in the English idiom “A picture is worth a thousand words” these complex ideas can be conveyed with just a single accurate and comprehensive image. In other words, an image of the molecular beacons in the dark phase base-pairing with a target and transforming into the hybrid light phase conveys the essence of its structural changes more effectively than a detailed description. Let us now examine further the factors which affect teaching and learning.

Factors which Influence Teaching and Learning

From the research in cognitive psychology, certain factors which influence teaching and learning have been identified, and are discussed in turn below.

Learning Produces Changes in the Brain

Learning promotes change. In other words, when we learn we change. Until recently, however, we did not have a clear idea about how learning actually brings about this change. As I mentioned in the literature review in Chapter Two, good teaching is often perceived as a craft or an art. Good teachers seem to be imbued with special skills that are hard to define (Zull, 2004). At the present time, however, with the advances made because of research in the fields of education, cognitive science, we are privy to deeper insights into what happens during the process of learning. Cognitive psychology teaches us that there are “fundamental neurological processes that happen in the brain as we learn” (p. 68). Learning produces physical change in the brain. Earlier models of the brain presented it as having a fixed structure, with the neuronal network that produced learning already in place. Research in the last few decades, however, has demonstrated that the fixed idea of the brain is erroneous, and has been replaced by a more “plastic” notion, in which the brain alters its own neuronal network almost continuously. As educators, especially science educators, we should be looking for new ideas of teaching supported by the growing understanding we have of the physical brain. Three such new factors are discussed below.

Do not just explain. Research in cognitive psychology has been advocating that knowledge consists of networks of neurons. One can extrapolate, therefore, that students' (i.e., novices') knowledge may be physically different from teachers' (i.e., experts') knowledge because students' neuron networks are different. If the students' network and the teacher's network are fundamentally different, Zull (2004) recommended that teachers should reduce their explanations and perform more demonstrations, and introduce more metaphors and recount more stories. This would promote more "show and tell," or more "showing" in the classroom rather than "explaining." In my study, the presenter of the teaching intervention used a narrative style to describe what Yaarasi et al. (2005) actually did, how things did go wrong at first, and how they learnt from their mistakes. These anecdotal accounts brought the scientific work to life, and impinged upon the impressionable minds of first-year undergraduates, encouraging them to become more responsible for their own learning. In addition, Zull (2004) claimed that explaining does not activate the emotions required for bringing about change in the neuronal networks.

Cognitive psychology research indicates that the positive emotions in learning are produced in those parts of the brain which are most heavily used when students develop their own ideas (Sylvester, 1995, p. 7). Such areas are embedded deep in the brain where they act as control centers for movements undertaken voluntarily. This research concludes that the biochemical rewards of learning are not achieved by providing explanations but by students taking ownership for their own learning.

Build on misconceptions/errors. As educators we are prone to treat pre-instructional misconceptions that students have as obstacles to learning. We should really search out, embrace, and welcome them, because these misconceptions can become the "raw materials for helping students build knowledge" (Zull, 2004, p. 71). For example, after an initial explanation of what a fair test is, the erroneous conceptions that students harbour of this concept which I have listed in Table 2 in Chapter Five, can be distributed to the students. Students could then be asked to explain why each of the items in this table is an erroneous conception, and how it is short of an accurate definition. Spotting the erroneous conceptions and addressing them in class is a very efficient teaching strategy according to Zull (2004). In this way, childhood, pre-instructional networks which are persistent, problematic, and resistant to change, as shown earlier in this chapter, can be used as a tool to enhance teaching, especially about procedural knowledge, and Concepts of Evidence in particular.

Engage the "whole brain." Another craft in the art of teaching is to develop techniques for engaging several regions of the brain simultaneously in the learning process. It is particularly important to involve different parts of the cerebral cortex, the part of the brain most responsible for its cognitive functions. A greatly simplified way to view the cerebral cortex is to divide into four regions (Zull, 2004): (1) the sensory cortex (for getting information or experience); (2) the integrative cortex near the sensory cortex (for making meaning and reflection); (3) the integrative cortex in the front (for creating new ideas and abstraction); and (4) the motor cortex (for acting on those ideas and active testing) (p. 72). As more regions of the cerebral cortex are used, more neural networks change, and thereby more learning occurs. Kolb (1984) claimed that deep learning is achieved by going through a sequence of *experience*, *reflection*, *abstraction* and *active testing* in that order. In my research activity the initial reading of the APL can be considered as the *experience*. Using the meanings included in the APL to understand this text, and then taking the pre-test lead to the *reflection*. New ideas are then created in the front, integrative cortex of the students as they watch the teaching intervention and this is the research *abstraction*. The final action is the *active testing* in the post-test. Therefore the lay-out and sequence of my research activity does have the potential to promote deep learning.

Earlier in this chapter, I discussed how descriptive drawings were used in this study to promote conceptual change in students. In this way, teaching interventions could target the correction of erroneous conceptions. As mentioned earlier, the dual coding theory proposed by Paivio (1990) would require two regions of the brain (i.e., cerebral cortex) – one for the verbal subsystem and the other for the imagery subsystem – to be stimulated simultaneously. This was achieved during the teaching intervention used in my study and would therefore have produced more changes in the connections that make up the neural networks resulting in more learning.

Teaching the Concepts of Evidence

An overwhelming majority of the instructors in my study were of the opinion that it is important to introduce students to these Concepts of Evidence and address them while teaching chemistry to first-year undergraduate students. There were several suggestions as to how to actually conduct the teaching of these concepts. Most instructors were of the opinion that they should teach the Concepts of Evidence piecemeal, with small portions addressed at one time, during each year of the chemistry undergraduate degree. The fourth-year students, in most post-secondary chemistry departments, have to do an investigative research project over two semesters. Some instructors were sure that the more complex Concepts of Evidence would be applicable and could be dealt with then. Many of the instructors felt that that they were teaching some of these Concepts of Evidence at present as well, though they may not be specifically labelling them as such. Some of the instructors voiced the opinion that teachers in high schools could perhaps first introduce the Concepts of Evidence when dealing with the different types of variables and the issue of sampling.

Student 38 was fully supportive of having the Concepts of Evidence taught explicitly because according to this student, "... it helped me to get an idea of why we were doing all these variables and controls and stuff" (p. 19). The explicit teaching about Concepts of Evidence has made Student 38 think more deeply about variables and perhaps consider how and when they are to be used. It has also made him look at controls, the purposes they serve, and how exactly they are to be set up. This deeper level of thinking was prompted by the questions asked in the pre- and post-tests and in the interviews. Asking probing questions and making perceptive comments have a huge role in fathoming the understanding of students especially during interviews.

Retention in Science

In the currently, fast-advancing fields of science and technology, the retention of students is an issue of national concern. According to O'Neal, Wright, Cook, Perorazio, and Purkiss (2007), approximately 40% of students who initially choose science majors, switch to other fields. Although this topic has been extensively researched, we still do not fully understand why this is the case. At present, various strategies are being employed to attract and retain undergraduates in science. One such attempt, at the University of Alberta, is Science 100, an interdisciplinary science course being offered to first-year undergraduates. Similarly, it is anticipated that teaching about evidence, and the Concepts of Evidence in particular, may help to interest and retain second-year students in science. This is because when these students are taught the Concepts of Evidence in one discipline of science such as chemistry, they will be able to notice the same threads and connections running through other disciplines as well. Such an endeavour will spark the imagination, excite the curiosity, foster creativity, and deepen understanding, thereby not only increasing the fascination for this subject, but hopefully promoting retention in science as well.

Conclusions

The main conclusions drawn from this study are summarized and discussed below.

Contributions to Research

In Chapter Three, I explained the conceptual frameworks on which I have based my study. Constructivism vouches for qualitative methods (Lincoln & Guba, 2000), and supports the belief that all individuals have their own ways of constructing knowledge. Learning occurs only when the learner actively participates in the process of knowledge construction. Each learner has to take the knowledge imparted by the teacher and personally reconstruct the meaning of this piece of knowledge (Millar, 1989): only then does learning really occur.

There is a chasm in perception between constructivism on one side and Gott's theoretical framework on the other. According to Gott's model of science, knowledge itself can be differentiated into two main types: substantive knowledge and procedural knowledge. Constructivism, on the other hand, treats knowledge as one entity. In Gott's model, facts contribute to substantive understanding, while skills furnish the procedural understanding. It is by the interaction of both substantive and procedural knowledge that cognitive processes are stimulated to solve problems. My work, which is steeped in constructivism, can be used as a bridge to span the gap between constructivism and Gott's model for science (1995).

From my study, it is evident that a certain level of substantive or content knowledge is required before it is possible to grasp the Concepts of Evidence and the role they play in the

APL. For example, from the instructors' interviews, one of the themes I arrived at was "The Thinking Gets kind of lost in the Content." Instructor 3 commented that some of these Concepts of Evidence could be taught very superficially but could also be addressed at a deeper level. Instructor 3 offered the Concept of Evidence known as "Reproducibility" as one such example. When delving deeply into reproducibility, one has to be very competent in dealing with the substantive knowledge represented in the APL. If one has to struggle with the substantive knowledge, it will be difficult to understand reproducibility at a deeper level. The deeper thinking about these Concepts of Evidence, therefore, gets immersed in the substantive knowledge.

One of the most-debated ideas that emerged from the analysis of the instructors' interviews was the "Status of substantive and procedural knowledge in the APL." The recurring issue that arose was that in order to understand and appreciate the Concepts of Evidence (i.e., procedural knowledge) embedded in this APL, a certain level of substantive knowledge is essential. These findings strongly support Gott's model of science, which promotes an interaction of substantive and procedural understanding to activate the cognitive processes and thereby solve problems

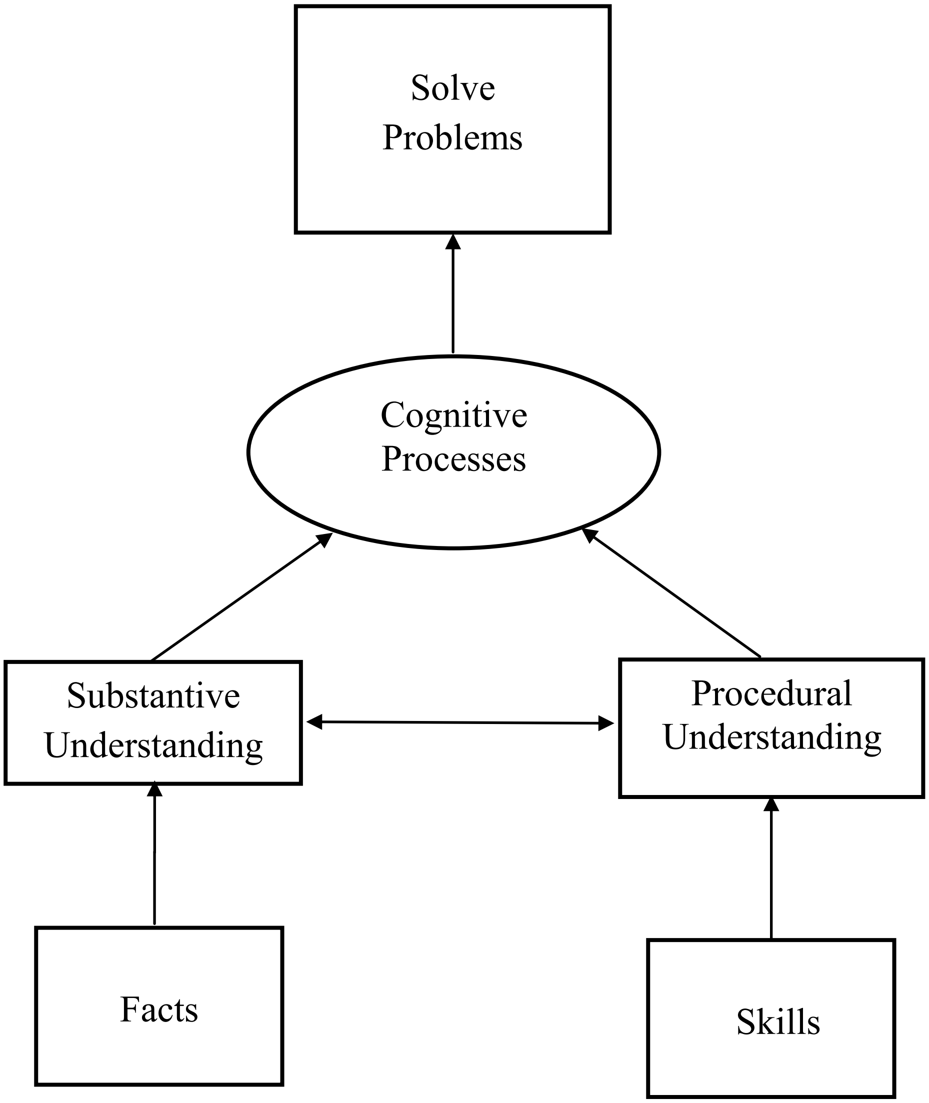
According to Gott et al.'s (1995) model for science, however, the arrow from substantive understanding to the cognitive processes is unidirectional. So too, is the arrow linking procedural understanding to the cognitive processes. In addition, there is no direct connection between substantive knowledge and procedural knowledge. Recent work from Gott and his colleagues (Roberts et al., 2010) sheds light on the importance of substantive and procedural understanding in performing open-ended science investigations. They took the stance that substantive and procedural understanding are "two connected but distinct types of knowledge" (p. 377). They emphasized the need at times to address procedural knowledge specifically and separately from substantive content (p. 377). My work also suggests that a certain degree of substantive knowledge is required for a proper understanding of procedural knowledge and that understanding procedural knowledge thoroughly promotes substantive knowledge as well. Therefore, I would modify Figure 2 in Chapter Two by inserting a bidirectional arrow connecting substantive knowledge and procedural knowledge. This would indicate that substantive knowledge promotes procedural understanding and vice versa. I have shown this modification clearly in Figure 8. My study seems to suggest that there is a mutually-enhancing, iterative relationship between substantive knowledge and procedural understanding.

To my knowledge, this is the first time that qualitative data have been collected on how first-year undergraduate students understand Concepts of Evidence. The role of qualitative data from the interviews is to uncover the "story" behind the data obtained from the evidence survey. Although the interview data are the leading data, the evidence survey data complement it, and both are needed to obtain the whole "story." I do not know of any other study where a qualitative approach like this has been used to study the understanding that first-year undergraduate students have of Concepts of Evidence, employing two methods of gathering the data - both interviews and an evidence survey. Another unique feature of this study is that it examines the views on, and the understanding of, the Concepts of Evidence held by both first-year undergraduate students doing chemistry courses and their instructors teaching them these courses. These are, therefore, two contributions to the research on Concepts of Evidence.

One of the greatest criticisms by science educators (G.P. Thomas, personal communication, June 29, 2012) levelled at the classification of Concepts of Evidence by Gott et al. (2003) is that there are too many categories and sub-categories. Unfortunately this has not

Figure 8

Modified Model for Science



(Based on Gott & Mashiter, 1991)

been documented or addressed in the science education literature and what I describe is an original contribution. I have adapted a scheme that designates each of these 21 categories of Concepts of Evidence to one of four types. This overall designation (as presented in Figure 9) reduces the 21 categories into four manageable groups. If one wants to consider "Measurement," for example, one can look up the seven Concepts of Evidence under that group and decide which one to choose. The access to, and the recall of, the appropriate Concept of Evidence are therefore made much easier, more systematic and less cumbersome.

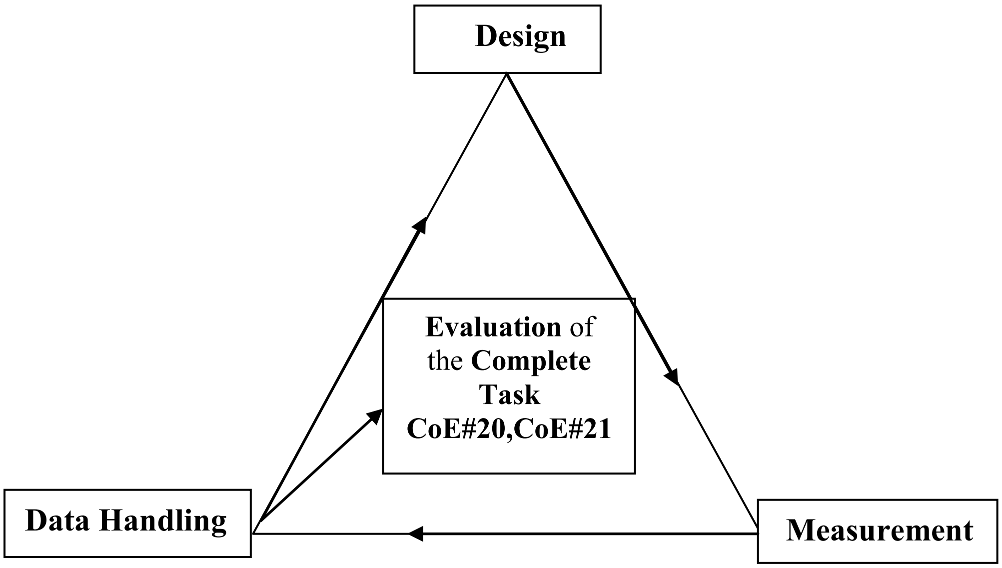
Gott et al.'s (1995) research about two decades ago contributed to the earlier model of science which only takes into consideration how individuals construct their own understanding about scientific concepts and solve problems in science. This initial model is based on a constructivist understanding of how learning occurs. Science in the 21st century, however, is a social practice and includes the social construction of scientific knowledge as well. This notion is steeped in constructivism which traces the spectrum of knowledge construction from the individual creation of knowledge in science to the social construction of scientific knowledge (Phillips, 1995).

Osborne's (2013) recent model takes into consideration three overarching practices of science: investigating; developing explanations and solutions; and evaluating. The social practice that Osborne termed "investigating" can be covered by the Concepts of Evidence in Gott's classification, which I have designated to the two groups "Design" and "Measurement." These two groups, as depicted in Figure 9, involve a total of 14 Concepts of Evidence. Similarly, Osborne's "developing explanations and solutions" is covered by the five Concepts of Evidence assigned to the group I have termed "data handling." The central "column" of Figure 3 called "evaluating" is covered by CoE#20 and CoE#21, which I have placed in the group "evaluation of the complete task." Therefore, all three practices of science are represented by particular groups of these Concepts of Evidence. In addition, all the 21 Concepts of Evidence as laid out by Gott et al. (2003) are included in the three practices of science stipulated by Osborne's (2013) model. It

Figure 9

21 Categories of Concepts of Evidence Designated to Four Groups

- Reliability & Validity of the Design (CoE #16)
- Design: Tables (CoE #15)
- Design: Accuracy and Precision (CoE #14)
- Design: Choosing Values(CoE #13)
- Design: Validity, “Fair Tests” and Controls (CoE # 12)
- Design of Investigations: Variable Structure (CoE #11)
- Fundamental Ideas (CoE #1)



- Statistical Treatment of Measurements Of a Single Datum (CoE #9)
- Reliability & Validity of a Datum (CoE #10)
- Data Presentation (CoE #17)
- Statistics for Analysis of Data (CoE #18)
- Patterns & Relationships in Data (CoE #19)

- Observation (CoE #2)
- Measurement (CoE #3)
- Instruments: Underlying Relationships (CoE#4)
- Instruments: Calibration Error (CoE #5)
- Reliability & Validity of a Single Measurement (CoE #6)
- The Choice of an Instrument for Measuring a Datum (CoE #7)
- Sampling a Datum (CoE #8)

<u>N.B.</u>

- Reliability & Validity of the Data in the Whole Investigation (CoE #20)
- Relevant Societal Aspects (CoE #21)

appears as though Gott's classification of Concepts of Evidence fits neatly into Osborne's model of the practices of science.

Osborne (2013) argued that in the 21st century where there is an "oversupply of information" (p. 266), the resource that is scarce is the inherent ability to make sense of all this information (Gilbert, 2005). In the next few decades, only societies that maintain their competitive edge will be classified as 'post-scientific' (Osborne, 2013, p. 266). The hallmarks of this 'post-scientific' society will be to highly value certain higher-order competencies such as "the ability to draw on a range of disciplinary knowledge; to think creatively and evaluate new ideas in a critical, reflective and rational manner; the ability to communicate and synthesize knowledge in an original manner; and the ability to solve problems creatively" (p. 266). Therefore, the competency and ability that will be most valued in the future will be higher order reasoning required for evaluation, synthesis, and critique. These valuable traits were promoted in first-year undergraduate students by my study.

Higher order reasoning in science requires knowledge of the concepts, facts, and theories of science (i.e., substantive knowledge). However, exploring 'how we know' and 'why we know' demands that students have a "knowledge of the standard procedures used by science and the epistemic constructs and criteria that guide practice" (i.e., procedural knowledge) (p. 273). There is ample evidence of the importance of procedural knowledge in scientific reasoning provided by many research studies (Millar, Lubben, Gott, & Duggan, 1995). If students are to develop expertise with scientific reasoning, it is imperative that they be given opportunities to analyse and interpret data and to construct and evaluate experimental designs (Osborne, 2013, p. 274). First-year undergraduate students were offered such an opportunity by participating in my study, because they were explicitly taught the Concepts of Evidence "as an essential feature of scientific inquiry" (p. 274) through the teaching intervention.

Moving Forward: Future Research Possibilities

Some of the most important questions that arise from my study are: (1) Is the conceptual change observed in my study permanent? (2) Or, does the conceptual change that has been achieved at present revert back to the original misconceptions over time? (3) If there is a reversal, what are the implications of this for pedagogy? I cannot answer these questions from my findings and results so far, because my post-test was done a week or two after the pre-test. However, by doing a post-test after one or two months, or even after a year, it would be possible to answer this question. I would have to do a more longitudinal study to answer these three questions adequately. There are three major longitudinal studies that could be recommended: (1) a five-year study tracing the development of the understanding of Concepts of Evidence held by the same cohort of students from Grade 10 to Grade 11 to Grade 12, and then in the first year and second year of their undergraduate degree. This long-term longitudinal study would shed light on the understanding of Concepts of Evidence held by these students as they transition from high school science to science at the university undergraduate level. (2) a four-year study following a cohort of students and their understanding of Concepts of Evidence of as they move from the first year to fourth year of their undergraduate science program. (3) It would also be worthwhile to similarly follow the understanding of Concepts of Evidence of pre-service science teachers as they go through their undergraduate years from Year 1 to Year 4.

Reflections on Research

I come from an Indian Orthodox Christian family where education is given high priority. My parents and two of my aunts are science teachers. My husband, daughter, parents, and I greatly value this opportunity that I have been granted to do a doctoral thesis in science education. It is only the faith that this opportunity was given to me by God that motivates me to press forward despite all obstacles.

As a researcher and educator, I have thoroughly enjoyed doing this research with people (i.e. students and instructors) as participants. I found it challenging but extremely fulfilling to collect my data from eight post-secondary institutions across Alberta. I had never been to most of these institutions and I was unknown to the students and instructors there. I think this helped tremendously because as a researcher, I was not too close to the participants or to the institutions, and I could easily maintain a professional distance from both. This kept my own biases in check, and to a minimum. It was also a challenge to share my research enthusiastically with my participants who were relative strangers.

This research activity was also unique because of its interdisciplinary nature. It was interdisciplinary on several counts, as shown below:

1. It was mainly a combination of science (i.e., chemistry) and science education.
2. The subject matter in the APL was interdisciplinary because it involved contributions from various science disciplines, such as photochemistry, photobiology, chemistry, biology, statistics, spectrophotometry, and fluorometry.
3. It catered to interdisciplinary science courses such as Science 100, offered at the University of Alberta.
4. It involved spectrophotometric and fluorometric analytical techniques because both absorbance and fluorescence intensity were measured.

The fact that this research project entails primarily a combination of science (i.e., chemistry) and science education brings multiple layers into the understanding of Concepts of Evidence. The chemistry part focussed on whether this discourse on Concepts of Evidence, employing the APL was chemically correct. In this case, the APL was only used as the vehicle for testing the students' understanding of these concepts. There would also be a concern about whether this research activity would contribute to a deeper understanding of chemistry. The science education part, however, put more emphasis on the actual pedagogy, how best to teach these Concepts of Evidence so that students have a deeper understanding which could be applied across other chemical contexts. Science education would be concerned more with what would be the best method to teach and test this understanding of Concepts of Evidence. This interdisciplinary nature, adding layers to the understanding of Concepts of Evidence, produces a discourse which is not entirely objective, but contributes to the culture and politics behind chemistry. By taking into consideration all the science disciplines, especially chemistry, science education, and the philosophy of science, which contribute to this project, it is possible to bring multiple lenses to bear on one's understanding of these Concepts of Evidence.

The main focus of my study is on students and their learning of these Concepts of Evidence, although I do get a perception of the instructors' views on these concepts and their pedagogical role, through interviewing them. Teaching these Concepts of Evidence initially in the first-year of the undergraduate degree would mean putting more responsibility on the instructors, who are researchers first and not pedagogues. One of the main reservations offered by instructors in teaching these Concepts of Evidence was the lack of time due to the fact that they have a tightly-packed curriculum to cover in first-year chemistry courses.

Recommendations

For the Educational Practice of Science

The message that I have from my research for chemists who are working with undergraduates would be that introducing and teaching these Concepts of Evidence "piecemeal," starting in the first year would be a very good way of exposing students to “the thinking behind the doing of science” (Gott & Duggan, 1995, p. 26). If this teaching was done in the first year and continued throughout the undergraduate years up to the fourth year, students would be well-equipped to plan and conduct their research projects which are mandatory in the final year of the undergraduate degree. Furthermore, if these Concepts of Evidence were taught and well understood in chemistry, students would recognize certain common “threads of understanding” running through all the disciplines of science. This would not only add to the undergraduates’ understanding of the “how to do” of science, but would excite their curiosity, capture their imagination, and result in greater engagement with science.

The teaching of these Concepts of Evidence could begin as early as junior high school, when students are introduced to variables. My message to teachers at this level is that students in this age category are very curious and eager to use the dexterity they have. This would be an ideal juncture, for example, to introduce the concept of “variables” with an activity. The teacher could then introduce Concept of Evidence #11, which deals with different types of variables. This could only be done if teachers, as pedagogues, know about Concepts of Evidence.

In teacher education, pre-service teachers need to be introduced to the Concepts of Evidence, if they have not met them earlier. A very convenient juncture for this introduction to procedural knowledge in science would be their “Science Curriculum and Instruction” classes and their “Pedagogical Content Knowledge for Science” classes. Pre-service teachers could be introduced to these Concepts of Evidence per se, and they could then be tested on their understanding of these concepts using APL. There are many other possibilities including the usage of e-Class to provide the APL as well as to deliver the pre- and post-tests.

For Research in Science Education

I have the responsibility to take what I have learned from this study and ensure that it is a part of the current thinking in science education. In this regard, I have written a paper, with my supervisor and a Canada Research Chair for Science Education, titled “The Role of Concepts of Evidence in the Teaching of Science.” This introductory paper is published in the journal most read by science teachers in Alberta, which is the “Alberta Science Education Journal.” This initial article introduces the reader to what these Concepts of Evidence are, and why they should be taught to students taking science courses in junior high and high school. This is the initial step that I have taken to disseminate my pedagogy messages to science teachers and educators.

There are a number of questions that arise out of my research such as: (1) Are students cognitively aware when they use these Concepts of Evidence as they handle data? (2) What sort of thinking are students to have when they evaluate data using these Concepts of Evidence? (3) Is there a difference in the understanding of Concepts of Evidence by undergraduates from different disciplines in the field of science? (4) What is the mechanism by which conceptual changes occur? (5) How can we promote conceptual change through better pedagogy? (6) Why does this resistance to conceptual change occur and what instructional methods can overcome it? (7) How might e-Class be used in a more effective constructivist pedagogical engagement?

For Curriculum Development

If the Concepts of Evidence can be introduced as early as Grade 9, then it is important that procedural knowledge is laid out as clearly in the science curriculum as the substantive knowledge is stipulated in it. Only if this is implemented will procedural knowledge be tested just as the substantive knowledge is examined. Compulsory testing may help to ensure that more teachers in the K-12 school system make the effort to specifically address Concepts of Evidence as a definite part of their teaching. This impetus to teach procedural knowledge appears to be lacking at present for teachers in the K-12 system.

Specifying the procedural knowledge in the curriculum, as clearly as the substantive knowledge is spelt out in this document, would have serious implications for the pedagogy of science in high school and undergraduate contexts. The topics mentioned in the curriculum could be taught more effectively by adopting a certain sequencing of the substantive knowledge and the procedural knowledge. The sequencing could be adjusted so that the substantive knowledge automatically leads into certain procedural knowledge (or Concepts of Evidence) or vice versa. In order to be effective, this would require science educators and curriculum experts to identify the points in the substantive content where the sequencing could be accommodated naturally. Implementing this sequencing will require a considerable amount of planning and preparation on the part of the teachers. They would have to find and choose appropriate activities to introduce these Concepts of Evidence such as doing experiments; carrying out investigations; watching demonstrations; or interpreting the data collected from participating in such empirical tasks.

In my study, I have used APL as the tool for testing the understanding of these Concepts of Evidence. The argumentative genre adopted by the APL makes it very similar to, but simpler than, primary literature. In addition, the APL preserves the canonical structure and the authentic results obtained by the researchers. This gives students the chance to see first-hand how data are handled by researchers, and how the evidence obtained from the data is coordinated with the theory. Besides, learning through APL may promote scientific ways of thinking in students. If teachers want to use APL for their teaching and testing of the Concepts of Evidence, they could use this adapted literature the way I did in my study, as a tool or vehicle to achieve their objective. Researchers in science education could then be encouraged to produce more APL from relevant, interesting and subject-specific primary literature.

Impact

I was struck by the “wow” factor expressed lucidly by one of my student participants who remarked enthusiastically and with conviction:

> I would never have even thought of the fact that you could make something glow, DNA glow. Just the fact they could figure out damage in base pairs through it glowing was just amazing. (S57, p. 35)

Some students really appreciated that this was an activity which was done for their benefit, to improve their understanding. They applauded the fact that it was not competitive or success-oriented. From this I learnt that students appreciate teaching and learning when it is done with their best interest in mind. I end with the encouraging remark of one student who had this to say about this research activity “It gave me something which will always be with me” (S47, p. 39).

Summary

This chapter provided a synopsis of the responses obtained to my research questions. The seminal research question which asked whether Concepts of Evidence can be taught or learnt was answered resoundingly in the affirmative with 86% of the student participants demonstrating an improvement in their performance on the post-test. This showed that pedagogy makes a significant difference in how students understand Concepts of Evidence. This understanding can be improved by using appropriate teaching interventions. It was evident from the scores on the pre-test that students did not enter this research activity as tabula rasa, but had their own pre-instructional conceptions which may be in line with or opposed to the scientific view. When these pre-instructional responses are opposed to the scientific view they are referred to as misconceptions or alternate conceptions.

The teaching intervention was praised by most students for its efficacy and its usefulness. They highly appreciated the auditory and visual stimulation, and helped them understand the substantive knowledge and the procedural knowledge. All the instructors were certain that the teaching intervention had a significant role in improving the students' understanding of the substantive knowledge and the procedural knowledge. The majority of instructors thought that including Concepts of Evidence when teaching chemistry to first-year undergraduate students was a very good and feasible idea. They stipulated that students should be taught about them "piecemeal" starting in the first year and extending to the final year. One instructor, however, was of the opinion that students should not be taught Concepts of Evidence but should be left to struggle with them on their own. Another instructor voiced the opinion that the primary literature should be chosen carefully to incorporate topics which would appeal to first-year undergraduate students. In my study, 65% of students liked the APL, while 4% found it boring.

The Concepts of Evidence were not directly addressed in the APL, but were introduced and explained in the teaching intervention. Since the Concepts of Evidence were embedded in the APL, it was essential for students to have a good understanding of the substantive knowledge in order to understand how the evidence concepts were applied. The substantive knowledge in the primary literature was interdisciplinary in nature, and that made the content challenging. It was this complexity which was simplified in the APL. Four themes emerged from the student interviews and five from the instructor interviews. These themes were related to each other as conveyed by the conceptual framework traced for each category. The positive aspects the students experienced by going through this research activity outweighed the negative experiences.

Reading is a constructive process in which the readers create meaning by forming inferential links between their background knowledge and the text they are reading. Literacy is constitutive of science itself. Therefore, if we desire our students to get insight into the nature and manner of scientific reasoning, we have to provide them with ample opportunities in using and exploring the language of science. The biggest barrier to students developing a deep understanding of what the texts they read mean is the vocabulary specific to these highly specialized fields.

In the APL that I created, I have minimized the barrier by providing a running list of the more complicated scientific words used in the primary literature. In my study, three-quarters of the students stated that they enjoyed reading the APL. Different sections of the APL evoked diverse types of challenges for the students. The argumentative genre adapted by the APL is highly conducive to developing scientific reasoning because it demonstrates what that looks like and sanctions that mode of thought. The APL conveys an uncertainty referred to as hedging which is characteristic of the scientific endeavor. The APL depicts science as a communal activity rather than an individual endeavor. Scientific knowledge is created by individuals interacting with one another in ways that modify their observations, theories and patterns of reasoning. These characteristics of the APL may make science less intimidating for the students.

From the constructivist viewpoint, there is a fundamental view of reading and a more critical view. The monolithic view is that reading is decoding and that is figuring out what the words are. Critical reading is not taught in high school or in university undergraduate level. Emphasis must be placed on scientific language which is easy to recall and useful for those who do not want to make a career in science, but want to remain scientifically literate citizens. The APL focuses on three aspects of scientific texts: (1) text structure; (2) text epistemology; and (3) meta-scientific language.

Students greatly appreciated the fact that the teaching intervention was on-line. The conversational style and the dual stimulation (i.e. by sight and sound) reinforced the learning.

The greater the sensory stimulation the greater is the learning. Personal anecdotes that the chemistry professor interspersed into the video made the teaching intervention more intimate and meaningful. The video lasted for only 32 minutes which is within the attention span of first-year undergraduates. Students could access the video at any time at their convenience. These features made the teaching intervention a very appropriate pedagogical tool. Only four Concepts of Evidence were covered in the teaching intervention. The professor who presented the teaching intervention referred to the Concepts of Evidence as the underlying formalism of how experiments are designed to display certain characteristics. To understand the Concepts of Evidence it was essential to have a fundamental understanding of the content knowledge. Conversely, a deep understanding of the Concepts of Evidence enhanced the substantive knowledge as well.

Scaffolding is provided by the teaching intervention which has tremendous efficacy in furthering the students' understanding of both the substantive knowledge and the procedural knowledge of the APL. Scaffolding, promoted by constructivism, is a style of instruction that provides students with the intellectual support to function at the cutting edge of their individual development. The presenter of the teaching intervention had not only content knowledge, but the pedagogical content knowledge and skill to communicate and understand the needs of the students. The teaching intervention was not test-score driven or curriculum-based. Learning was designed to be for fun, so that the goals of deriving meaning and understanding which are implicit to constructivism were met.

Teachers have to build on the existing neuronal network of their students. In my study, this initial neuronal network is tapped through the pre-tests. Knowledge resides in the brain as neuronal networks. These neuronal networks are the physical equivalent of knowledge, and the changes in the connections which make up the network constitute learning. Repeated usage of neuronal networks, lead to the branching of neurons and the formation of interconnections between neurons which increases the density of that region of the brain.

Vygotsky (1978) stated that cognitive development stems from social interaction, especially those providing opportunities for guided learning. During the interview each student was given the opportunity to interact with “a more knowledgeable other.” The interview which was scaffolded effectively has the following traits: (1) common goal or intersubjectivity; (2) ongoing diagnosis and adaptive support; (3) dialogues and interactions; and (4) fading and transfer of responsibility. It is the cognitive effort of the individual that results in the construction of knowledge. Not all students have the same understanding of substantive knowledge and procedural knowledge. The degree to which students are prepared to “wrestle” with the new content knowledge or the newly-introduced Concepts of Evidence, decides their level of understanding. This is in full agreement with constructivism which does not endorse the “spectator theory” of learning. According to constructivism the learner has to be an active participant or actor in the learning process, and only then does deep learning occur.

Vygotsky (1978) also focussed on the role of culture in the development of higher mental functions. The zone of proximal development (ZPD) is “the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance, or in collaboration with more capable peers” (p. 86). ZPD signifies the range of tasks that a learner can perform with the help and guidance of others but cannot yet perform independently. This zone represents the region in which this transfer of ability from the shared environment to the individual occurs. ZPD has two development levels: (1) actual development level; and (2) the level of potential development. In my study, the actual development level is the initial level of understanding of APL and Concepts of Evidence attained by the students independently. The level of potential development is the upper limit of tasks that one can perform with the help of a more competent individual. This level of potential development in my study was demonstrated by the post-test of the evidence survey. According to Vygotsky, the ZPD is the area where the most sensitive instruction or guidance should occur. When I interviewed the students, I had the opportunity to promote the growth in the understanding of individual students within their zone of proximal development. The extent of the reconstruction of knowledge depends on the degree of the interrelationship between the APL and the teaching intervention. If the interrelationship is dynamic, it will promote the individual reconstruction of pre-existing knowledge.

The ground-breaking work in the field of conceptual change was carried out by Posner, Strike, Hewson and Gertzag (1982). Learning is achieved through the interaction between new knowledge and one's existing preliminary conceptions, and the final outcome of learning is determined by the nature of this interrelationship. The conceptions that people have which are different from the scientifically correct ones are termed misconceptions or alternative conceptions etc. It is therefore vital to have concept-change-oriented instruction geared towards addressing some of these erroneous conceptions, so that the thinking of students who hold these unscientific views can undergo conceptual change, and eventually lead to cognitive development. When the misconceptions students have about certain concepts are known, the teaching intervention could be tailored to address them. This would give students who possess these erroneous conceptions a chance to face and rectify them because these alternate conceptions hinder the development of a deep understanding. As new information is assimilated, new nodes will be formed and nodes will be interconnected, resulting in webs of interconnected knowledge. When the teaching intervention is able to give the knowledge in the APL an overall structure (i.e. sequential or hierarchical), it will aid in the understanding, recall and retention of this information (Chi, 2005). Consequently, the information in the APL will be slotted into the appropriate positions in the new mental models formed by the students, promoting the understanding, the learning and the recall of this knowledge: this process in its entirety brings about a conceptual change.

According to Posner et al. (1982), four conditions have to be met for conceptual change to occur in an individual's understanding: (1) There must be dissatisfaction with existing conceptions; (2) The new conceptions offered must be intelligible; (3) The new conception must be initially plausible; and (4) The new conception must offer the possibility of a fruitful research program. In this study students who have displayed a conceptual change have a heightened level of comprehension. The process of achieving a conceptual change is not inherently difficult, but can prove to be challenging when students are not aware of their misconceptions and they lack the alternative ontologically distinct categories to which they should shift their misconceptions (Chi & Roscoe, 2002). The task of conceptual change requires effort on the part of the learner. Pre-existing knowledge plays a crucial role in attaining conceptual change. Conceptual change is a gradual process. There are four schools of thought among researchers in the field of conceptual change. Descriptive drawings can be used very effectively to promote conceptual change.

Research in cognitive psychology suggests that there are certain important factors which influence teaching and learning. They are: (1) Learning produces changes in the brain; (2) When teaching do not just explain; (3) Build on misconceptions as you teach; and (4) Engage the whole brain when you teach. The conclusions that can be drawn from this study were then summarized in this final chapter.

REFERENCES

Abelson, P. H. (1980). Scientific communications. *Science, 209,* 60-63.

Aboud, F. (1988). *Children and prejudice*. Oxford, England: Blackwell.

Achinstein, P. (1983). Concepts of evidence. In P. Achinstein (Ed.), *The concept of evidence.* Oxford, England: The Oxford University Press.

Achinstein, P. (2001). *The book of evidence.* Oxford, England: Oxford University Press.

Aikenhead, G. S. (2004). Science-based occupations and the science curriculum: Concepts of Evidence. *Science Education, 89*, 242-275.

Aikenhead, G. S. (2006). *Science education for everyday life: Evidence-based practice.* New York, NY: Teachers College Press.

Aikenhead, G. S., & Elliott, D. (2010). An emerging decolonizing science education in Canada. *Canadian Journal of Science, Mathematics and Technology Education, 10*, 321-338.

Aikenhead, G. S., & Michell, H. (2011). *Bridging cultures: Indigenous and scientific ways of knowing nature.* Toronto, Ontario: Pearson Canada Inc.

Aikenhead, G. S., & Ogawa, M. (2007). Indigenous knowledge and science revisited. *Cultural Studies of Science Education, 2*, 539-620.

Akkaya, C. G. (2003). *Comparison of teaching the relation rate course using the traditional and the experimental training methods on students success in lycee 2 classes.* (Unpublished master'a thesis). Marmara University, Istanbul.

Alberts, B. (2005). A wakeup call for science faculty. *Cell, 123*, 739-741.

Alexander, D. J. (1974). *Nuffield Secondary Science: An evaluation.* London, England: Macmillan.

Alexander, P. A. & Jetton, T. L. (2000). Learning from text: A multidimensional and developmental perspective. In M.L. Kamil, P.B. Mosenthal, P.D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, p. 285-310). New Mahwah, NJ: Lawrence Erlbaum.

Allie, S., Buffler, A., Kaunda, L., Campbell, B, & Lubben, F. (1998). First-year physics students' perceptions of the quality of experimental measurements. *International Journal of Science Education, 20,* 447-459.

American Association for the Advancement of Science (1967). *Science – A process approach.* Washington, DC: Ginn & Co.

Anderson, C. W., & Smith, E. L. (1984). Children's preconceptions and content-area textbooks. In G.G. Duffy, L.R. Roehler, & J. Mason (Eds.), *Comprehension instruction: Perspectives and suggestions* (pp. 187-201). New York, NY: Longman.

Anderson, T. R. (2007). Bridging the educational research-teaching practice gap. *Biochemistry and Molecular Biology Education, 35*(6), 465-470.

Andre, T., Whigham, M., Hendrickson, A., & Chambers, S. (1999). Competency beliefs, gender affect, and gender stereotypes of elementary students and their parents about science versus other school subjects. *Journal of Research in Science Teaching, 36,* 719-747.

Appleyard, B. (1992). *Understanding the present*. London, England: Picador.

Armstrong, H. E. (1896). How science must be studied to be useful. In G. V. Praagh (Ed.), *H. E. Armstrong and science education: Selections from the teaching of scientific method and other papers on education* (pp. 41-50). London, England: John Murray.

Ary, D., Jacobs, L. C., Razavieh, A., & Sorenson, C. (2007). *Introduction to research in education* (7th ed.). Belmont, CA: Wadsworth.

Arzi, H. J., Ben-Zvi, R. & Ganiel, U. (1984). Can teachers speak for their students? A comparison between teachers' and students' evaluation of a school science course. *European Journal of Science Education, 6,* 379-86.

Audi, R. (Ed.). (1999). *The Cambridge dictionary of philosophy.* Cambridge, England: Cambridge University Press.

Ausubel, D. P. (1963). *The psychology of meaningful verbal learning.* New York, NY: Grune & Stratton, Inc.

Ausubel, D. P., Novak, J. D., & Hanesian, H. (1978). *Educational psychology: A cognitive view* (2nd ed.). New York, NY: Holt, Rinehart and Winston.

Babbie, E. (1995). *The practice of social research* (7th ed.). Belmont, CA: Wadsworth.

Babco, E. L. (1994). Women in engineering. *Comments, 32*(4), 22.

Bae, Y., Choy, S., Geddes, C., Sable, J., & Snyder, T. (2000). *"Trends in educational equity of girls and women"* [NCES 2000-030, U. S. Department of Education, National Center for Education Statistics (NCES), Washington, DC, 2000].

Baker, L., Dreher, M. J., & Guthrie, J. T. (Eds.). (2000). *Engaging young readers.* New York, NY: Guildford.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Banerjee, A. C. (1995). Teaching chemical equilibrium and thermodynamics in undergraduate general chemistry classes. *Journal of Chemical Education, 72*(10), 879-881.

Baram-Tsabari, A., & Yarden, A. (2005). Text genre as a factor in the formation of scientific literacy. *Journal of Research in Science Teaching, 42*, 403-428.

Barefoot, B. O. (2000). The firs-year experience: Are we making it any better? *About Campus/January-February* 2000, 12-18.

Barrows, H. S. (1980). *Problem-based learning: An approach to medical education*. New York, NY: Springer Publishing.

Battiste, M. (2000). *Reclaiming Indigenous voice and vision.* Vancouver, BC: UBC Press.

Battiste, M. (2002). *Indigenous knowledge and pedagogy in First Nations education: A literature review with recommendations.* Ottawa, ON: Indian and Northern Affairs Canada.

Becker, G. S. (1975). *Human capital: A theoretical and empirical analysis, with special reference to education*. New York, NY: NBER.

Belenky, M. F., Clinchy, B. M., Goldberger, N. R., & Tarule, J. M. (1986). *Women's ways of knowing: The development of self, voice, and mind*. New York, NY: Basic Books.

Bencze, L., & Hodson, D. (1999). Changing practice by changing practice: Toward more authentic science and science curriculum development. *Journal of Research in Science Teaching, 36*(5), 521-539.

Bennett, J. (2003). *Teaching and learning science.* London, England: Continuum Books.

Berelson, B. (1952). *Content Analysis in communication research*. New York, NY: The Free Press.

Berger, K. S. (2009). *The developing person through childhood and adolescence.* New York: Worth Publishers.

Berger, P. L., & Luckman, T. (1966). *The social construction of reality: A treatise in the sociology of knowledge*. New York, NY: Anchor.

Berk, L. E. & Winsler, A. (1995). *Scaffolding children's learning: Vygotsky and early childhood education*. Washington, DC: National Association for the Education of Young Children.

Birchard, K. (2001). Arts and science want you! *Chronicle of Higher Education 47*:A6.

Bird, A. (1998). *Philosophy of science.* Montreal, Canada: McGill-Queen's University Press.

Blankenship, J. C. (1991). *Attrition among male nursing students.* Unpublished doctoral dissertation, Department of Adult Education, University of Georgia, Athens.

Bodner, G. (1986). Constructivism: A theory of knowledge. *Journal of Chemical Education, 63*, 873-878.

Bordt, M., De Broucker, P., Read, C., Harris, S., & Zhang, Y. (2001). Determinants of science and technology skills: Overview of the study. *Education Quarterly Review, 8*(1), 8-12.

Bourdieu, P. (1973). Cultural reproduction and social reproduction. In R. Brown (Ed.), *Knowledge, education and cultural change: Papers in the sociology of education.* London, England: Tavistock Publications Limited.

Bourgeois, E., Duke, C., Guyot, J.-L., & Merrill, B. (1999). *The adult university*. Milton Keynes, England: Open University Press.

Bowden, M., & Doughney, J. (2010). Socio-economic status, cultural diversity and the aspirations of secondary students in the Westourern Suburbs of Melbourne, Australia. *Higher Education, 59*, 115-129.

Bowden, M., & Doughney, J. (2012). The importance of cultural and economic influences behind the decision to attend higher education. *The Journal of Socio-economics, 41*, 95-103.

Bowlby, J., & McMullen, K. (2002). *At a crossroads: First results from the 18 to 20-year old cohort of the Youth in Transition Survey*. Ottawa, ON: Human Resources Development Canada and Statistics Canada.

Bowles, S., Gintis, H., & Osborne, M. (2001). The determinants of earnings: A behavioral approach. *Journal of Economic Literature XXXIX*, 137-1176.

Bozkoyun, Y. (2004). *Facilitating conceptual change in learning rate of reaction concepts.* (Unpublished master's thesis). Middle East Technical University, Ankara.

Bracht, G. H., & Glass, G. V. (1968). The external validity of experiments. *American Educational Research Journal, 5*, 437-74.

Brannen, J. (2005). *Mixed methods research: A discussion paper* (Economic and Social Research Council National Centre for Research Methods Review Papers). Retrieved from www.ncrm.ac.uk/research/outputd/publications/documents/MethodsReviewPaperNCRM-005.pdf

Bransford, J., Sherwood, R., Vye, N & Rieser, J. (1986). Teaching thinking and problem-solving. *American Psychologist, 41*, 1078-1089.

Brewer, J., & Hunter, A. (1989). *Multimethod research: A synthesis of styles.* Newbury Park, CA: Sage.

Brewer, M. B., & Collins, B. E. (Eds.). (1981). *Scientific inquiry and the social sciences* (a volume in honor of Donald T. Campbell). San Francisco, CA: Jossey-Bass.

Brewer, W. F., & Lambert, B. L. (1993). The theory-ladenness of observation: Evidence from cognitive psychology. In *Proceedings of the fifteenth annual conference of the cognitive science society* (pp. 254-259). Hillsdale, NJ: Erlbaum.

Brickhouse, N. W., Dagher, Z. R., Shipman, H. L., & Letts, W. J. (2002). Evidence and warrants for belief in a college astronomy course. *Science and Education, 11,* 573-588.

Brickhouse, N. W., & Kittleson, J. M. (2006). Visions of curriculum, community and science. *Educational Theory, 56*(2), 191-204.

Brill, G., & Yarden, A. (2003). Learning biology through research papers: A stimulus for question-asking by high school students. *Cell Biology Education, 2,* 266-274.

Brill, G., Falk, H., & Yarden, A. (2004). The learning process of two high-school biology students when reading primary literature. *International Journal of Science Education, 26*(4), 497-512.

Brotman, J. S., & Moore, F. M. (2008). Girls and science: A review of four themes in the science education literature. *Journal of Research in Science Teaching, 45*, 971-1002.

Brown, J., Collins, A., & Duiguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*, 32-42.

Brownlow, S., Jacobi, T., & Rogers, M. (2000). Science anxiety as a function of gender and experience. *Sex Roles, 42*, 119-131.

Bruner, J. (1996). *The culture of education.* London, England: Harvard University Press.

Bryman, A. (2007). Barriers to integrating quantitative and qualitative research. *Journal of Mixed Methods Research, 1*(1), 8-22.

Buffler, A., Allie, S., & Lubben, F. (2001). The development of first year physics students' ideas about measurement in terms of point and set paradigms. *International Journal of Science Education, 23*, 1137-56.

Buffler, A., Lubben, F., & Ibrahim, B. (2007). *The relationship between introductory physics students' views on the nature of science and the nature of scientific measurement*. Paper presented at the 6th international conference of the European Science Education Research Association (ESERA), August 21-25, in Malmö, Sweden.

Bunce, D. M., Flens, E. A., & Neiles, K. Y. (2010). How Long Can Students Pay Attention in Class? A Study of Student Attention Decline Using Clickers. *Journal of Chemical Education, 87* (12), 1438–1443.

Cajete, G. A. (1999). *Igniting the sparkle: An indigenous science education model.* Skyand, NC: Kivaki Press.

Cakir, M. (2008). Constructivist approaches to learning in science and their implications for science pedagogy: A literature review. *International Journal of Environmental & Science Education, 3*(4), 193-206.

Çalik, M., & Ayas, A. (2005). A comparison of level of understanding of grade 8 students and science student teachers related to selected chemistry concepts. *Journal of Research in Science Teaching, 42*(6), 638-667.

Çalik, M., Kolomuç, A., & Karagölge, Z. (2010). The effect of conceptual change pedagogy on students' conceptions of rate of reaction. *Journal of Science Educational Technology, 19*, 422-433.

Çalik, M., Ünal, S., Coştu, B., & Karataş, F. Ö. (2008). Trends in Turkish science education. *Essays in Education*, Special Edition, 23-45.

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), *Handbook of research on teaching*. Chicago, IL: Rand McNally.

Canadian Association of University Teachers (2014). *The CAUT almanac of postsecondary education in Canada (2013-2014)* p. 23, 27.

Canadian Council on Learning (2007). Gender differences in career choices: Why girls don't like science. *Lessons in Learning, November 1,* 1-11.

Canadian Council on Learning (2007b). *The cultural divide in science education for aboriginal learners.* Retrieved from: http://www.cclcca.ca/CCL/Reports/LessonsInLearning/LinL20070116_Ab_sci_edu.htm

Carey, S. (1985). *Conceptual change in childhood.* Cambridge, MA: MIT Press.

Carey, S. (1986). Cognitive science and science education. *American Psychologist, 41,* 1123-1130.

Carnegie Melon University. (1996). *Best Practices for Teaching First-year Undergraduates*: Retrieved from http://www.cmu.edu/teaching/documents/bestpractices.htm

Carney, T. F. (1971). Content analysis: A review essay. *Historical Methods Newsletter, 4*(2), 52.

Carter, A. O., Griffin, G. H., & Carter, T. P. (2006). A survey identified publication bias in the secondary literature. *Journal of Clinical Epidemiology, 59*, 241-245.

Cartwright, N. (1983). *How the laws of physics lie.* Oxford, England: Clarendon Press.

Cassidy, J. C., & Johnson, R. E. (2002). Cognitive test anxiety and academic performance. *Contemporary Educational Psychology, 27,* 270-295.

Çetin, P. S., Kaya, E., & Geban, O. (2009). Facilitating conceptual change in gases concepts. *Journal of Science Education Technology, 18,* 130-137. doi:10.1007/s10956-008-9138-y

Champagne, A. B., Gunstone, R. F., & Klopfer, L. E. (1985). Instructional consequences of students' knowledge about physical phenomena. In L. H. T. West & A. L. Pines (Eds.), *Cognitive structure and conceptual change* (pp. 61-90). Orlando, FL: Academic Press.

Chalmers, A. F. (1982). *What is this thing called science?* (2nd ed). Milton Keynes, England: Open University Press.

Chi, M. T. H. (2000a). Cognitive understanding levels. In A. E. Kazdin (Ed.), *Encyclopedia of Psychology, Vol. 2*, (pp. 172-175). Oxford: Oxford University Press.

Chi, M. T. H. (2000b). Self-explaining: The dual processes of generating inference and repairing mental models. In R. Glaser (Ed.), *Advances in Instructional Psychology: Educational Design and Cognitive Science, Vol. 5.* Lawrence Erlbaum Associates Publishers.

Chi, M. T. H. (2005). Common sense conceptions of emergent processes: Why some misconceptions are robust. *Journal of Learning Sciences, 14*, 161-199.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121-152.

Chi, M. T. H., & Roscoe, R. D. (2002). The processes and challenges of conceptual change. In M. Limón & L. Mason (Eds.), *Reconsidering Conceptual Change. Issues in Theory and Practice* (pp. 3-27). Netherlands: Kluwer Academic Publishers.

Chi, M. T., Roscoe, R. D., Slotta, J. D., Roy, M., & Chase, C. C. (2012). Misconceived causal explanations for emergent processes, *Cognitive Science, 36*, 1-61.

Chiappetta, E. L., Fillman, D. A., & Sethna, G. H. (1991). A method to quantify major themes of scientific literacy in science textbooks, *Journal of Research in Science Teaching, 28*(8), 713-725.

Chiappetta, E. L., Sethna, G. H., & Fillman, D. A. (1993). Do middle school life science textbooks provide a balance of scientific literacy themes? *Journal of Research in Science Teaching, 30*(7), 787-797.

Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research, 63*, 1-49.

Chinn, C. A., & Brewer, W. F. (2000). Knowledge change in response to data in science, religion and magic. In K. S. Rosengren, C. N. Johnson, & P. L. Harris (Eds.), *Imagining the* *impossible: Magical, scientific, and religious thinking in children* (pp. 334-371). Cambridge, England: Cambridge University Press.

Chinn, C. A., & Brewer, W. F. (2001). Models of data: A theory of how people evaluate data. *Cognition and Instruction,19*, 323-393.

Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education, 86,* 175-218.

Coffey, H. (2009). *Zone of proximal development*. Learn NC. K-12 Teaching and Learning from the UNC School of Education. Retrieved from: http://www.learnnc.org/lp/pages/5075

Cohen, S. (1984). Justification and truth. *Philosophical Studies, 46*, 279-295.

Coll, R., & Treagust, D. (2001). Learners' use of analogy and alternative conceptions for chemical bonding: A cross-age study. *Australian Science Teaching Journal, 48*(1), 24-32.

Collingwood, R. G. (1956). *The idea of history.* Oxford, England: Oxford University Press.

Collins, H. M. (1985). *Changing order*. London, England: Sage.

Confrey, J. (1990). A review of the research on student conceptions in mathematics, science, and programming. In C. G. Cazden (Ed.), *Review of research in education, 16*, 3-56. Washington, DC: American Educational Research Association.

Considine, G., & Zappala, G. (2002). The influence of social and economic disadvantage in the academic performance of school students in Australia. *Journal of Sociology, 38*, 129-148.

Cook, T. D. & Campbell, D. T. (1979). *Quasi-experimentation: design & analysis issues for field settings.* Chicago, IL: Rand McNally College Publishing Company.

Çostu, B., & Ayas, A. (2005). Evaporation in different liquids: Secondary students' conceptions. *Research in Science and Technology Education, 23*, 73-95.

Côté, J. E. & Levine, C. (2002). *Identity formation, agency, and culture: A social psychological synthesis.* Mahwah, NJ: Lawrence Erlbaum Associates.

Courtenay, B. C., Merriam, S. B., & Reeves, P. M. (1998). The centrality of meaning-making in transformational learning: How HIV-positive adults make sense of their lives. *Adult Education Quarterly, 48*(2), 65-84.

Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches* (2nd ed.). Thousand Oaks, CA: Sage.

Creswell, J. (2008). *Educational research; Planning, conducting, and evaluating quantitative and qualitative research.* Upper Saddle River, NJ: Pearson Education, Inc.

Crocker, J., & Major, B. (1989). Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review, 96,* 608-630.

Crossan, B., Field, J., Gallacher, J., & Merrill, B. (2003). Understanding participation in learning for non-traditional adult learners: Learning careers and the construction of learning identities. *British Journal of Sociology of Education, 24*, 55-67.

Dagher, Z. R. & BouJaoude, S. (1997). Scientific views and religious beliefs of college students: The case of biological evolution. *Journal of Research in Science Teaching, 34,* 429-445.

Davis, J. (2001). Conceptual Change. In M. Orey (Ed.), *Emerging perspectives on learning, teaching, and technology*. Retrieved Sept. 8, 2015, from http://projects.coe.uga.edu/epltt/

Davis, B., Sumara, D., & Luce-Kapler, R. (2000). *Engaging minds: Learning and teaching in a complex world.* London, England: Lawrence Erlbaum Associates.

de Jong, T., de Hoog, R., & Schreiber, G. (1988). Knowledge acquisition for an integrated project management system. *Information Processing and Management, 24,* 681-691.

de Jong, T., & Ferguson-Hessler, M. G. M. (1996). Types and qualities of knowledge. *Educational Psychologist, 31*(2), 105-113.

De Regts, H. W., & Dieks, D. (2003). *Contextual approach to scientific understanding.* Retrieved from http://philosciarchive.pitt.edu/archive/00001354/

Denzin, N. K. (1978). *The research act: A theoretical introduction to sociological method* (2nd ed.). New York, NY: McGraw-Hill.

Denzin, N. K., & Lincoln, Y. S. (Eds.). (1998). *The landscape of qualitative research: Theories and issues.* London, England: Sage.

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2008). *The landscape of qualitative research: Theories and issues.* London, England: Sage.

Dewey, J. (1938). *Experience and education.* New York, NY: Collier and Kappa Delta Pi.

Dewey, J. (1960). *The quest for certainty.* New York, NY: Capricorn.

Dewey, J. (1969). *The school and society* (reprinted as a joint edition with *The child and the curriculum*). Chicago, IL: University of Chicago Press.

Diamond, M. C. (1988). *Enriching heredity: The impact of the environment on the anatomy of the brain.* New York, NY: Free Press.

DiGisi, L. L., & Willett, J. B. (1995). What high school biology teachers say about their textbook use: A descriptive study. *Journal of Research in Science Teaching, 32*(2), 123-142.

diSessa, A. A. (2002). Why conceptual ecology is a good idea. In M. Limon & L. Mason (Eds.), *Reconsidering conceptual change: Issues in theory and practice* (pp. 29-60). Dordrecht: Kluwer.

Drew, C. (2011). Why science majors change their minds (It's just so darn hard). *The New York Times,* November 4.

Driver, R. (1981). Pupils' alternate frameworks in science. *European Journal of Science Education, 3*, 93-101.

Driver, R., Asoko, H., Leach, J., Mortimer, E., & Scott, P. (1994). Constructing scientific knowledge in the classroom. *Educational Researcher, 23*(7), 5-12.

Driver, R., & Easley, J. (1978). Pupils and paradigms: a review of literature related to concept development in adolescent science student. *Studies in Science Education, 5,* 61-84. doi:10.1080/03057267808559857

Driver, R., Guesne, E., & Tiberghien, A. (Eds.). (1985). *Children's ideas in science.* Milton Keynes, England: Open University Press.

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

Driver, R., & Oldham, V. (1985). A constructivist approach to curriculum development. *Studies in Science Education, 13,* 105-122.

Driver, R., Squires, A., Rushworth, P., & Wood-Robinson, V. (1994). *Making sense of secondary science: Research into children's ideas*. London, England: Routledge.

Dudczak, C.A. (1994). *An Inverse Inference model of evidence.* Paper presented at the Annual Meeting of the Eastern Communication Association (Washington, DC).

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

Duggan, S., Johnson, P., & Gott, R. (1996). A critical point in investigative work: Defining variables. *Journal of Research in Science Teaching, 33*, 461-474.

Duit, R., & Treagust, D. F. (1998). Learning in science – From behaviourism towards social constructivism and beyond. In B. J. Fraser & K. G. Tobin (Eds.), *International handbook of science education* (Vol. 1, pp. 3-25). Dordrecht, Netherlands: Kluwer Academic.

Dunbar, K. (1995). How scientists really reason: Scientific reasoning in real-world laboratories. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 365-397). Cambridge, MA: MIT Press.

Durant, J.R. (1994). What is scientific literacy? *European Review, 2*(1), 83-89.

Durant, J. R., Evans, G. A., & Thomas, G. P. (1992). Public understanding of science in Britain: The role of medicine in the popular presentation of science. *Public Understanding of Science 1*, 161-182.

Duschl, R. A. (1990). *Restructuring science education.* New York, NY: Teachers College Press.

Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (Eds.). (2006). *Taking science to school: Learning and teaching science in Grades K-8* (Committee on Science Learning, Kindergarten through Eighth Grade, Board of Science Education, Center for Education, Division of Behavioural and Social Sciences and Education). Washington, DC: The National Academic Press.

Eberlein, T., Kampmeier, J., Minderhout, V., Moog, R. S., Platt, T., Varma-Nelson, P., & White, H. B. (2008). Pedagogies of engagement in science: A comparison of PBL, POGIL and PLTL. *Biochemistry and Molecular Biology Education, 36*(4), 262-273.

Edens, K. M., & Potter, E. (2003). Using descriptive drawings as a conceptual change strategy in elementary science. *School Science and Mathematics, 103*(3), 135-144.

Edwards, R. (1993). *Mature women students: Separating or connecting family and education*. London, England: Taylor & Francis.

Egan, S. K., & Perry, D. G. (2001). Gender identity: A multidimensional analysis with implications for psychosocial adjustment. *Developmental Psychology, 37,* 451-463. doi:10.1037/0012-1649.37.4.451

Ehninger, D., & Brockreide, W. (1978). *Decision by debate* (2nd ed.). New York, NY: Harper & Row.

Engineers Canada. (2014). *Canadian engineers for tomorrow: Trends in engineering enrolment and degrees awarded 2009-2013*. Ottawa, Canada: Prism Economics and Analysis.

Epstein, H. T. (1970). *A strategy for education.* Oxford, England: Oxford University Press.

Ertmer, P. A., & Ottenbreit-Leftwich, A. T. (2010). Teacher technology change: How knowledge, confidence, beliefs and culture intersect. *Journal of Research on Technology in Education, 42*(3), 255-284.

Falk, H., Brill, G., & Yarden, A. (2005). Scaffolding learning through research articles by a multimedia curriculum guide. In M. Ergazaki, J. Lewis, & V. Zogza (Eds.), *Proceedings of The Vth conference of The European Researchers in Didactics of Biology (ERIDOB)* (pp. 175-192). Patras, Greece.

Falk, H., Brill, G., & Yarden, A. (2008). Teaching a biotechnology curriculum based on adapted primary literature. *International Journal of Science Education, 30*(14), 1841-1866.

Falk, H., & Yarden, A. (2009). "Here the scientists explain what I said." Coordination practices elicited during the enactment of the Results and Discussion sections of adapted primary literature. *Research in Science Education*, Published Online. doi:10.1007/s11165-008-9114-9

Fensham, P. J. (1992a). Science education at first degree level. *International Journal of Science Education, 14*(5), 505-514.

Fensham, P. J. (1992b). Science and technology. In P. W. Jackson (Ed.). *Handbook of research on curriculum* (pp.789-829). New York, NY: Macmillan.

Fensham, P. J., Gunstone, R. F., & White, R. T. (1994). *The content of science: A constructivist approach to its teaching and learning.* Washington, DC: The Falmer Press.

Feyerabend, P. (1975). *Against method*. London, England: Verso.

Fincham, F., Kashdan, T., & Rose, P. (2004). Curiosity and exploration: Facilitating positive subjective experiences and personal growth opportunities. *Journal of Personality Assessment, 82*, 31-34.

Fine, G. (2003). Introduction. *Plato on knowledge and forms: Selected essays* (p. 5). Oxford, England: Oxford University Press.

Flynn, A. B. (2012). A post-class question strategy that provides feedback and connects in- and out-of-class learning. *Collected Essays on Teaching and Learning, 5*, 153-160.

Flynn, A. B., & Amellal, D. G. (2016). Chemical information literacy: pK_a values – Where do students go wrong? *Journal of Chemical Education, 93*, 39-45.

Flynn, A. B., Caron, J., Laroche, J., Daviau-Duguay, M., Marcoux, C., & Richard, G. (2014). Nomenclature 101.com: A free student-driven organic chemistry nomenclature learning tool. *Journal of Chemical Education, 91*, 1855-1859.

Foley, P. (2007). *The socio-economic status of vocational education and training students in Australia*. In: N.C.f.V.E. (Ed.). Research Australian Government, Adelaide, Australia.

Ford, D. J. (2009). Promises and challenges for the use of adapted primary literature in science curricula: Commentary. *Research in Science Education, 39,* 385-390.

Forrester, D., & Jantzie, N. (n. d.). *Learning Theories.* Retrieved from *itu.dk/~metteott/ITU_stud/Speciale/Lœring/Learning%20Theories.doc*

Freeman, C. E. (2004). *Trends in educational equity of girls and women: 2004* (NCES 2005-016, U.S. Department of Education, NCES, Washington, DC, 2004).

Freeman, M., de Marrais, K., Preissle, J., Roulston, K., & St. Pierre, E. A. (2007). Standards of evidence in qualitative research: An incitement to discourse. *Educational Researcher, 36*(1), 25-32.

Frost, S., Reiss, M., & Frost, J. (2005). Count me in! Gender and minority ethnic attainment in school science. *School Science Review, 86 (316)*, 105-112.

Gagne, R. M. (1965). *The psychological basis of science – a process approach*, AAAS miscellaneous publication, 65-68. Washington, DC: AAAS.

Gall, M. D., Borg, W. R., & Gall, J. P. (1996). *Educational research: An introduction.* (6th edn.) White Plains, NY: Longman Publishing.

Gee, B., & Clackson, S. G. (1992). The origin of practical work in the English school science curriculum. *School Science Review, 73*(265), 79-83.

Geertz, C. (1973). *The interpretation of cultures: Selected essays.* New York, NY: Basic Books.

Germann, P. J., & Aram, R. J. (1996). Students' performances on the science processes of recording data, analyzing data, drawing conclusions and providing evidence. *Journal of Research in Science Teaching, 33*, 773-798.

Gilbert, J. (2005). *Catching the knowledge wave? The knowledge society and the future of education.* Wellington, New Zealand: NZCER Press

Gilbert, J. K., & Boulter, C. J. (1991). *Developing models in science education*. Dordrecht, Netherlands: Kluwer.

Gilborn, D., & Gipps, C. (1996). *Recent research on the achievements of ethnic minority pupils,* Report for the Office for Standards in Education. London, England: HMSO.

Given, L. M. (Ed.) (2008). *The Sage encyclopedia of qualitative research methods.* Los Angeles, CA: Sage.

Glaser, R. (1984). Education and thinking: The role of knowledge. *American Psychologist, 39*, 93-104.

Glasersfeld, E. V. (1987). Learning as a constructive activity. In C. Janvier (Ed.), *Problems of representation in the teaching of mathematics* (pp. 3-17). Hillsdale, NJ: Lawrence Erlbaum.

Glasersfeld, E. V. (1993). Questions and answers about radical constructivism. In K. Tobin (Eds.), *The practice of constructivism in science education* (pp. 23-38). Washington, DC: AAAS Press.

Glesne, C. (1998). *Becoming qualitative researchers: An introduction*. Ontario, Canada: Longman.

Glynn, S.M., & Muth, D. (1994). Reading and writing to learn science: Achieving scientific literacy. *Journal of Research in Science Teaching, 31*(9), 1057-1073.

Godfrey-Smith, P. (2003). *Theory and reality: An introduction to the philosophy of science.* Chicago, IL: The University of Chicago Press.

Goodman, N., & Elgin, C. Z. (1988). *Reconceptions in philosophy and other arts and sciences*. London, UK: Routledge.

Gorard, S. (2003). *Quantitative methods in social science*. London, England: Continuum Press.

Gorard, S., & Taylor, C. (2004). *Combining methods in educational and social research*. Buckingham, England: Open University.

Gorlewski, J. A., Gorlewski, D. A., & Ramming, T. M. (2012). *Theory into practice: Case stories for school leaders.* New York, USA: Springer.

Goss, L. S. & Bernstein, D. A. (2005). *Teaching psychology: A step by step guide*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Gott, R., & Duggan, S. (1995). *Investigative work in the science curriculum.* Buckingham, England: Open University Press.

Gott, R., Duggan, S., & Roberts, R. (2003). *Concepts of evidence*. Durham, NC: Durham University. Retrieved 19/06/2009 from http://www.dur.ac.uk/rosalyn.roberts/Evidence/cofev.htm

Gott, R., Duggan, S., Roberts, R., & Hussain, A. (2003). *Research into understanding scientific evidence*. Retrieved from www.dur.ac.uk/richard.gott/Evidence/cofev.htm

Gott, R., & Johnson, P. (1999). Science in schools: time to pause for thought? *School Science Review*, *81(295),* 21-28.

Gott, R., & Mashiter, J. (1991). Practical work in science – a task-based approach? In B. E. Woolnough (Ed.) *Practical Science* (pp 53-67). Buckingham, PA: Open University Press.

Gott, R., & Murphy, P. (1987). *Assessing investigations at ages 13 and 15*. APU Science Report for Teachers No. 9. London, England. DES.

Gott, S. P. (1989). Apprenticeship instruction for real-world tasks: The coordination of procedures, mental models and strategies. *Review of Research in Education, 15,* 97-169.

Grace, A. (2010). *Space matters: Lifelong learning, sexual minorities and realities of adult education as social education*. Standing Conference on University Teaching and Research in the Education of Adults, University of Warwick, UK.

Graesser, A., Golding, J. M., & Long, D. L. (1991). Narrative representation and comprehension. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.). *Handbook of reading research* (Vol. 2, pp. 171-205). New York, NY: Longman.

Gredler, M. E. (1997). *Learning and instruction: Theory into practice.* NJ: Prentice Hall.

Green, K. C. (1989). A profile of undergraduates in the sciences. *The American Scientist, 78*, 475-480.

Griffiths, A. K., & Bettle, A. C. (1985). High school students' understanding of food webs. Identification of a learning hierarchy and related misconceptions. *Journal of Research in Science Teaching, 22*, 421-436.

Gruber, H. (1973). Courage and cognitive growth in children and scientists. In M. Schwebel & J. Raph (Eds.), *Piaget in the classroom* (pp. 73-105). New York, NY: Basic Books.

Guba, E. G., & Lincoln, Y. S. (1998). Competing paradigms in qualitative research. In N. K. Denzin &Y. S. Lincoln (Eds.), *The landscape of qualitative research: Theories and issues* (pp. 221-259). London, England: Sage.

Guimond, S., & Roussel, L. (2001). Bragging about one's school grades: Gender stereotyping and students'perception of their abilities in science, mathematics, and language. *Social Psychology of Education, 4,* 275-293.

Gulliksen, H. (1950). *Theory of mental tests.* New York, NY: John Wiley.

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest, 8,* 1-51. doi: 10.1111/j.1529-1006.2007.00032.x

Hand, B., Alvermann, D., Gee, J., Guzzetti, B., Norris, S., Phillips, L., et al. (2003). Message from the "Island Group": What is literacy in science literacy? *Journal of Research in Science Teaching, 40*, 607-615.

Hanson, N. R. (1958). *Patterns of discovery: An inquiry into the conceptual foundations of science.* Cambridge, England: Cambridge University Press.

Hapgood, S., Magnusson, S. J., & Palinscar, A. S. (2004). Teacher, text, and experience: A case of young children's scientific inquiry. *The Journal of the Learning Sciences, 13*(4), 455-505.

Harding, S. (1993). Rethinking standpoint epistemology: "What is strong objectivity?" In L. Alcoff & E. Potter (Eds.), *Feminist epistemologies* (pp. 49-82). New York, NY: Routledge.

Harré, R. (1984). *The philosophies of science: an introductory survey* (2nd ed.). Oxford, England: Oxford University Press.

Harré, R. (1986). *Varieties of realism: A rationale for the natural sciences.* Oxford, England: Basil Blackwell

Harsh, J. A., Maltese, A. V., & Tai, R. H. (2012). Undergraduate research experiences from a longitudinal perspective. *Journal of College Science Teaching, 41*(1), 84-91.

Hatton, J., & Plouffe, P. B. (1997). *Science and its ways of knowing*. Upper Saddle River, NJ: Prentice Hall.

Henriques, L. (2002). Children's ideas about weather: A review of the literature. *School Science and Mathematics, 102*(5), 202-215.

Herman, C. (1999). Reading the literature in the jargon-intensive field of molecular genetics. *Journal of College Science Teaching, 28*(4), 252-253.

Hesse-Biber, S. N. (2010). *Mixed methods research: Merging theory with practice*. London, England: The Guilford Press.

Hewson, M. G., & Hewson, P. W. (1983). Effect of instruction using students'prior knowledge and conceptual change strategies on science learning. *Journal of Research in Science Teaching, 20*(8), 731-743. doi. 10.1002/tea

Hewson, M. G. (2015). *Embracing indigenous knowledge in science and medical teaching.* New York: Springer.

Hewson, P. W. (1992). *Conceptual change in science teaching and teacher education.* National Center of Educational Research, Documentation and Assessment, Madrid, Spain.

Hill, K. T. (1984). Debilitating motivation and testing: A major educational program, possible solutions, and policy applications. In R. E. Ames & C. Ames (Eds.), *Research on Motivation in Education* (Vol. 1, pp. 245-274). New York, NY: Academic Press.

Hinman, R. L. (1998). Who is scientifically literate anyway? *Phi Delta Kappan, 79*(7), 540-544.

Hodson, D. (1990). A critical look at practical work in school science. *School Science Review, 70 (256),* 33-40.

Hodson, D. (1996). Practical work in school science: Exploring some directions for change. *International Journal of Science Education, 18*(7), 755-760.

Hodson, D. (1998). *Teaching and learning science: Towards a personalized approach.* Buckingham, PA: Open University Press.

Hofer, B. K. (2004). Epistemological understanding as a metacognitive process. *Educational Psychologist, 39*(1), 43-55.

Hogan, D. M., Tudge, J (1999). Implications of Vygotsky's theory for peer learning. In A. M. O'Donnell & A. King (Eds.), *Cognitive perspectives on peer learning* (pp. 39-65). Mahwah, NJ: Erlbaum.

Holliday, W. G., Yore, L. D., & Alvermann, D. E. (1994). The reading-science learning-writing connection: Breakthroughs, barriers, and promises. *Journal of Research in Science Teaching, 31*(9), 877-893.

Horspool, W. M., & Song, P. (1995). *CRC Hand book of Organic Photochemistry and Photobiology*. Boca Raton FL: CRC Press.

Hoskins, S. G., Stevens, L. M., & Nehm, R. S. (2007). Selective use of the primary literature transforms the classroom into a virtual laboratory. *Genetics, 176,* 1381-1389.

Houde, A. (2000). Student symposia on primary research articles. *Journal of College Science Teaching, 30*, 184-187.

Howe, A. C. (1993). A Vygotskian perspective on teaching for conceptual change. In *The Proceedings of the Third International Seminar on Misconceptions and EducationalStrategies in Science and Mathematics,* Misconceptions Trust: Ithaca, NY.

Hug, B., & McNeill, K. L. (2008). Use of first-hand and second-hand data in science: Does data type influence classroom conversations. *International Journal of Science Education, 30*(13), 1725-1751.

Hurd, P. D. (2000). Science education for the 21st century. *School Science and Mathematics, 100(6)*, 282-288.

Hwang, W. S., Ryu, Y. J., Park, J. H., Park, E. S., Lee, E. G., Koo, J. M. ... Moon, S. Y. (2004). Evidence of a pluripotent human embryonic stem cell line derived from a cloned blastocyst. *Science, 303*, 1669-1674.

Hyland, K. (1996). Writing without conviction? Hedging in scientific research articles. *Applied Linguistics 17*(4): 433-454.

Ivarsson, J., Schoultz, J., & Saljo, R. (2002). Map reading versus mind reading: Revisiting children's understanding of the shape of the earth. In M. Limon & L. Mason (Eds.), *Reconsidering conceptual change: Issues in theory and practice* (pp. 77-99). Dordrecht:Kluwer.

James, R. (Ed.). (2002). *Socioeconomic background and higher education participation: An analysis of school students' aspirations and expectations.* Melbourne, Australia: Centre for the Study of Higher Education, University of Melbourne, Australia.

Janick-Buckner, D. (1997). Getting undergraduates to critically read and discuss primary literature. *Journal of College Science Teaching, 27*, 29-32.

Jarman, R., & McClune, B. (2002). A survey of the use of newspapers in science instruction by secondary teachers in Northern Ireland. *International Journal of Science Education, 24*(10), 997-1020.

Jiménez-Aleixandre, M. P., & Federico-Agraso, M. (2009). Justification and persuasion about cloning: Arguments in Hwang's paper and journalistic reported versions. *Research in Science Education.* doi:10.1007/s11165-008-9113-x.

Johnson, R. B. & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7). 14-26.

Johnson, R. B., & Turner, L. (2003). Data collection strategies in mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 297-320). Thousand Oaks, CA: Sage.

Johnson, S. (1987). Gender differences in science. *International Journal of Science Education, 9*, 467-481.

Johnston, M. (1999). Influences of adult attachment on exploration. *Psychological Reports, 84,* 31-34.

Jones, M. G., Howe, A., & Rua, M. J. (2000). Gender differences in students' experiences, interests and attitudes toward science and scientists. *Science Education, 84*, 180-192.

Joyce, J. M. (2004). *Williamson on evidence and knowledge*. Ann Arbor, Michigan: Blackwell Publishing Limited.

Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Boston, MA: Allyn and Bacon.

Kahle, J. B. (1982). The disadvantaged majority: Biology education for women and minorities. *The American Biology Teacher, 44*(6), 351-356.

Kahle, J., & Lakes, M. (1983). The myth of equality in science classrooms. *Journal of Research in Science Teaching, 20*, 131-140.

Kanari, Z. & Millar, R. (2004). Reasoning from data: How students collect and interpret data in science investigations. *Journal of Research in Science Teaching, 41*(7), 748-769.

Karmiloff-Smith, A. (1984). Children's problem solving. In M. Lamb, A. Brown, & B. Rogoff (Eds.), *Advances in developmental psychology* (Vol. 3, pp. 39-90). Hillsdale, NJ: Erlbaum.

Karmiloff-Smith, A. (1986). From meta-processes to conscious access: Evidence from children's metalinguistic and repair data. *Cognition, 23*, 95-147.

Karmiloff-Smith, A. (1988). The child is a theoretician, not an inductivist. *Mind and Language, 3*, 1-13.

Kerlinger, F. N. (1973). *Foundations of behavioural research* (2nd ed.). New York, NY: Holt, Rinehart & Winston.

Kimmelberg, H., & Norenberg, M. (April 1989). Astrocytes. *Scientific American,* 66-76.

Kind, P., Jones, K., & Barmby, P. (2007). Developing attitudes toward science measures. *International Journal of Science Education, 29*(7), 871-893.

Kish, L. (1965). *Survey sampling*. New York, NY: Wiley.

Klahr, D. & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science, 12*, 1-48.

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice Hall.

Korpan, C. A., Bisanz, G. L., Bisanz, J., Henderson, J. M. (1997). Assessing literacy in science: Evaluation of scientific news briefs. *Science Education, 81*(5), 515-532.

Koslowski, B. (1996). *Theory and evidence: The development of scientific reasoning.* Cambridge, MA: MIT Press.

Kotte, D. (1992). *Gender differences in science achievement in 10 countries.* Frankfurt, Germany: Peter Lang.

Koulaidis, V., & Ogborn, J. (1995). Science teachers' philosophical assumptions: How well do we understand them? *International Journal of Science Education, 17*, 273-282.

Krahn, H., & Lowe, G. S. (2002). (4th ed.). *Work, Industry and Canadian Society*. Toronto, ON: Thomson Nelson.

Krahn, H., & Taylor, A. (2005). Resilient teenagers: Explaining the high educational aspirations of visible-minority youth in Canada. *Journal of International Migration and Integration, 6*(3/4), 405-434.

Krippendorff, K. (2004) (2011). *Content analysis: An introduction to its methodology*. London, England: Sage.

Kroger, J. (2004). *Identity in adolescence: The balance between self and other.* London, UK: Routledge.

Krupa, M., Selman, R., & Jacquette, D. (1985). The development of science explanations in children and adolescents: A structural approach. In S. Chipman, J. Segal, & R. Glaser (Eds.), *Thinking and learning skills* (Vol. 2, pp. 427-455). Hillsdale, NJ: Erlbaum.

Kuh, G. D., Cruce, T. M., Shoup, R., Kinzie, J., & Gonyea, R. M. (2008). Unmasking the effects of student engagement on first-year college grades and persistence. *The Journal of Higher Education, 79*(5), 540-563.

Kuh, G. D., Kinzie, J., Buckley, J., Bridges, B., & Hayek, J. C. (2007). *Piecing together the student success puzzle: Research, propositions, and recommendations.* ASHE Higher Education Report, 32(5). San Francisco, CA: Jossey-Bass.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review, 96*(4), 674-689.

Kuhn, D. (1991). *The skills of argument*. Cambridge, MA: Cambridge University Press.

Kuhn, T. (1962). *The structure of scientific revolutions*. Chicago, IL: University of Chicago press.

Kuhn, T. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago, IL: University of Chicago press.

Kuhn, D., Amsel, E. & O'Loughlin, M. (1988). *The development of scientific thinking skills.* London, England: Academic Press.

Kuldell, N. (2003). Read like a scientist to write like a scientist. *Journal of College Science Teaching, 33*, 32-35.

Kumar, D. D. (2003). Trends in postsecondary science in the United States. Annals of the *American Academy of Political and Social Science, 585*, 124-133.

Kumar, D. D., & D. E. Chubin. (Eds.) (2000). *Science, technology, & society: A sourcebook on research and practice.* New York, NY: Kluwer Academic/Plenum.

Kumar, D. D., & Morris, J. D. (2005). Predicting scientific understanding of prospective elementary teachers: Role of gender, education level, courses in science, and attitudes toward science and mathematics. *Journal of Science Education and Technology, 14*(4), 387-391.

Lang, K., & Ruud, P. A. (1986). Returns to schooling, implicit discount rates and black-white wage differentials. *Review of Economics and Statistics, 68*, 41-47.

Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life.* Cambridge, MA: Cambridge University Press.

Layton, D. (1973). *Science for the people.* London, England: George Allen and Unwin.

Layton, D., Jenkins, E., Macgill, S. & Davey, A. (1993). *Inarticulate science? Perspectives on the public understanding of science and some implications for science education.* Driffield, East Yorkshire: Studies in Education Ltd.

Leaper, C., Farkas, T., & Brown, C. S. (2012). Adolescent girls' experiences and gender-related beliefs in relation to their motivation in Math/Science and English. *Journal of Youth Adolescence, 41,* 268-282.

Lemke, J. (1990). *Talking science: language, learning and values.* Norwood, NJ: Ablex.

Lentz, M., & Evans, R. (2005). *The extent to which primary sources in the biology classroom are a tool for teaching scientific literacy*. Winston-Salem, NC: Wake Forest University Press.

Lerner, R. (1993). A developmental contextual view of human development. In S. C. Hayes, L. J. Hayes, H. W. Reese, & T. R. Sarbin (Eds.), *Varieties of scientific contextualism*. Reno, NV: Context Press.

Levine, E. (2001). Reading your way to scientific literacy: Interpreting scientific articles through small-group discussions. *Journal of College Science Teaching, 31*(2), 122-125.

Limon, M. (2002). Conceptual change in history. In M. Limón & L. Mason (Eds.), *Reconsidering conceptual change: Issues in theory and practice,* 259-289. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry.* Beverly Hills, CA: Sage.

Lincoln, Y. S., & Guba, E. G. (2000). Paradigmatic controversies, contradictions and emerging confluences. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 163-188). Thousand Oaks, CA: Sage.

Lindlof, T. (1995). *Qualitative communication research methods*. Thousand Oaks, CA: Sage.

Locke, E. A. (1977). An empirical study of lecture notetaking among college students. *Journal of Educational Research, 77*, 93.

Loevinger, J. (1976). *Ego development: Conception and theory*. San Francisco, CA: Jossey-Bass.

Longino, H. (1993). Subjects, power and knowledge: Description and prescription in feminist philosophies of science. In L. Alcoff & E. Potter (Eds.), *Feminist epistemologies* (pp. 101-120). New York, NY: Routledge.

Lubben, F., & Millar, R. (1996). Children's ideas about the reliability of experimental data. *International Journal of Science Education, 18,* 955-968.

Lubben, F., Campbell, B., Buffler, A., & Allie, S. (2001). Point and set reasoning in practical science measurement by entering university freshmen. *Science Education, 85*, 311-327.

MacIvor, M. (1995). Redefining science education for Aboriginal Students. In M. Battiste & J. Barman (Eds.), *First Nations education in Canada: The circle unfolds* (pp. 73-98). Vancouver, BC: University of British Columbia Press.

Mackie, D., & Smith, E. (1998). Intergroup relations: Insights from a theoretically integrative approach. *Psychological Review, 105*, 499-529.

Madill, H. M., Campbell, R. G., Cullen, D. M., Armour, M-A., Einsiedel, A. A., Ciccocioppo, A-L., Sherman, J. ... Coffin, W. L. (2007). Developing career commitment in STEM-relatedfields: Myth versus reality. In R. J. Burke & M. C. Mattis (Eds.), *Women and minorities in science, technology, engineering and mathematics: Upping the numbers*. Northampton MA: Edward Elgar Publishing Inc.

Magnus, D., & Cho, M. K. (2005). Issues in oocyte donation for stem cell research. *Science, 308,* 1747-1748.

Mallow, J. V. (1986). *Science anxiety: Fear of Science and how to overcome it.* (Revised edition). Clearwater, FL: H & H Publications.

Mallow, J. V. (1994). Gender-related science anxiety: A first binational study. *Journal of Science Educational Technology, 3*(4), 227-238.

Markel, N. (1998). *Semiotic psychology: Speech as an index of emotions and attitudes*. New York, NY: Peter Lang.

Marker, M. (2004). The four r's revisited: Some reflections on first nations and higher education. In L. Andres, & F. Finlay (Eds.), *Student affairs – Experiencing higher education* (pp. 171-188). Vancouver, BC: UBC Press.

Mar-Molinero, C. (1991). Gender differences in undergraduate university admissions in the United Kingdom. *Educational Studies, 17*(1), 49-64.

Marshall, J. S. (2003). A veteran's view of science education today. *Review of Policy Research, 20,* 629-634.

Marshall, C., & Rossman, G. B. (2006). *Designing qualitative research* (4th ed.). Thousand Oaks, CA: Sage.

Martin, M. (1972). *Concepts of science education: A philosophical analysis.* London, England: Scott, Foresman and Company.

Masnick, A. M. & Morris, B. J. (2002). Reasoning from data: The effect of sample size and variability on children's and adults' conclusions. In W. D. Gray & C. D. Schunn (Eds.), *Proceedings of the 24th annual conference of the Cognitive Science Society* (p. 643-648). Mahwah, NJ: Lawrence Erlbaum.

Matthews, M. (1995). *Constructivism and New Zealand science education*. Auckland, New Zealand: Dunmore Press.

Matthews, M. (Ed.). (1998). *Constructivism in science education*. Dordrecht, Netherlands: Kluwer.

Mathews, M. R. (Ed.). (1989). *The scientific background to modern philosophy*. Indianapolis, IN: Hackett Publishing Company.

Mathews, M. R. (1993). Constructivism and science education: Some epistemological problems. *Journal of Science Education and Technology, 2*(1), 359-369.

Mathews, M. R. (1994). *Science teaching: The role of history and philosophy of science.* New York, NY: Routledge.

Mautner, T. (Ed.). (1999). *The Penguin dictionary of philosophy.* London, England: Penguin Books.

Mayer, R.E. (1983). What we have learned about increasing the meaningfulness of science prose? *Science Education, 67*(2), 223-237

McGrayne, S. B. (2001). *Nobel prize women in science: Their lives, struggles and momentous discoveries.* Washington, DC: National Academic Press.

McNeal, A. P., & C. D'Avanzo (Eds.) (1997). *Student-active science: Models of innovation in college science teaching.* Fort Worth, TX: Saunders College Publishing.

Meece, J. L., & Holt, K. (1993). A pattern analysis of students' achievement goals. *Journal of Educational Psychology, 85*, 582-590.

Mehra, B. (2002). Bias in qualitative research: Voices from an online classroom. *The Qualitative Report, 7(1*). Retrieved from http://www.nova.edu.login.ezproxy.library.ualberta.ca/ssss/QR/QR7-1/mehra.html

Merriam, S. B. (1998). Case studies as qualitative research. In *Qualitative research and case study applications in education*. (Rev. and expanded ed.; pp. 26-43). San Francisco, CA: Jossey-Bass.

Merrill, B. (1999). *Gender, change and identity: Mature women students in universities*. Aldershot, Hampshire, UK: Ashgate.

Mertl, M. (2000). Grants for women in science. *Science Careers,* Retrieved from http://sciencecareers.sciencemag.org/career_magazine/previous_issues/articles/2000_05_31/noDOI.16094295288747995254

Miles, M., & Huberman, M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.

Millar, R. (1985). Training the mind: Continuity and change in the rhetoric of school science. *Journal of Curriculum Studies, 17*(4), 369-382.

Millar, R. (1989). Constructive criticisms. *International Journal of Science Education, 11*, 587-596.

Millar, R. (1991). A means to an end: The role of processes in science education. In B. E. Woolnough (Ed). *Practical science: The role and reality of practical work in school science* (pp. 43-53). Philadelphia, PA: Open University Press.

Millar, R. (1994). What is “scientific method” and can it be taught? In R. Levinson (Ed.). *Teaching Science.* London, England: Routledge.

Millar, R. (1996). Towards a science curriculum for public understanding. *School Science Review*, 77(280), 7-18.

Millar, R., & Driver, R. (1987). Beyond processes. *Studies in Science Education, 14,* 33-62.

Millar, R., Leach, J., & Osborne, J. (Eds.). (2000). *Improving science education: The contribution of research.* Buckingham, PA: Open University Press.

Millar, R., Lubben, F., Gott, R., & Duggan, S. (1995). Investigating in the school science laboratory: Conceptual and procedural knowledge and their influence on performance. *Research Papers in Education, 9*(2), 207-248.

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

Miller, M. M., Boone, J. & Fowler, D. (1992, November). *The emergence of greenhouse effect on the issue agenda: A news stream analysis.* Paper presented at the annual meeting of the Midwest Association for Public Opinion Research, Chicago, IL.

Miller, S., & Fredericks, M. (2003). The nature of “Evidence” in qualitative research methods. *International Journal of Qualitative Methods, 2*(1). Article 4. Retrieved from http://www.ualberta.ca/~iiqm/backissues/2_1/html/miller.html

Moll, L. (1990). *Vygotsky and education. Instructional implications and applications of sociohistorical psychology.* Cambridge, England: Cambridge University Press.

Monk, M. & Osborne, J. (Eds.) (2000). *Good practice in science teaching: What research has to say.* Buckingham, PA: Open University Press.

Moore, E. (2000a). The changing patterns of university studies: Towards lifelong learning in Finnish universities. *Higher Education Management, 12*, 113-127.

Moore, E. (2000b). *Aikuisena yliopistossa: Yliopisto-opiskelijoiden ikärakenne ja 30 vuotta täyttäneiden opiskelijoiden elämänkulku* [University studies at adult age: The age structure of university students and the life course of students aged 30+; in Finnish]. Joensuu, Finland: Joensuu University Press.

Moore, E. (2003). *Pitkä opintie: Aikuisällä suoritettu yliopistotutkinto ja koulutuksellisen elämänkulun muutos* [Long study paths: Studying for a degree at adult and change in educational life course; in Finnish]. University of Joensuu Publications in Social Sciences, No. 61. Joensuu, Finland: Joensuun yliopistopaino.

Moore, E. (2006). Educational identities of adult university graduates. *Scandinavian Journal of Educational Research, 50*(2), 149-163.

Moreira, M. A. (1980). A non-traditional approach to the evaluation of laboratory instruction in general physics courses. *European Journal of Science Education, 2*, 441-448.

Muench, S. B. (2000). Choosing primary literature in biology to achieve specific educational goals. *Journal of College Science Teaching, 29*, 255-260.

Mulkay, M. (1979). *Science and the sociology of knowledge*. London, England: George Allen & Unwin.

Myers, G. (1992). Textbooks and the sociology of scientific knowledge. *English for Specific Purposes, 11,* 3-17.

Nakhleh, M. B. (1992). Why some students don't learn chemistry. *Journal of Chemical Education, 69*(3), 191-196.

Nash, R. (2005). Cognitive habitus and collective intelligence: Concepts for the explanation of inequaltiy of educational opportunity. *Journal of Educational Policy, 20*, 3-21.

National Research Council. (1995). *National Science convocation on undergraduate education.* Arlington, VA: National Research Council Center for Science, Mathematics and Engineering Education and Training.

National Research Council. (2000). *How people learn: Brain, mind, experience and school.* Washington, DC: National Academy Press.

National Science Foundation. (1970). *Science education: The task ahead for the National Science Foundation* (No. 71-13). Washington, DC: National Science Foundation.

National Science Foundation. (1994). *Women, minorities, and persons with disabilities in science and engineering: 1994.* (NSF 94-333). Arlington, VA: Science Foundation. Retrieved from http://www.nsf.gov/sbe/srs/nsf00327/access/toc.htm

National Science Foundation. (2006). *Research on gender in science and engineering FY 2006 (GSE)pProgram announcement,* NSF 05-614; Washington, DC: National Science Foundation.

National Science Foundation. (2010). *Women, minorities, and persons with disabilities in science and engineering: 2009.* (NSF 09-305); Washington, DC: National Science Foundation.

Neuendorf, K. A. (2002). *The content analysis guidebook*. Thousand Oaks, CA: Sage.

Newell, W. H. (2007). Distinctive challenges of library-based interdisciplinary research and writing: A guide. *Issues in Integrative Studies, 25,* 84-110.

Newman, D. (2005). Ego development and ethnic identity formation in rural American Indian adolescents. *Child Development, 76*, 734-746.

Niaz, M., Aguilera, D., Maza, A., & Liendo, G. (2002). Arguments, contradictions, resistances and conceptual change in students' understanding of atomic structure. *Science Education, 86*, 505-525.

Nicoll, G. (2001). A report of undergraduates' bonding misconception. *International Journal of Science Education, 23*, 707-730.

Norris, S. P. (1984). Defining observational competence. *Science Education, 68*, 129-142.

Norris, S. P. (1985). The philosophical basis of observation in science and science education. *Journal of Research in Science Teaching, 22*(9), 817-833.

Norris, S. P. (1992). Practical reasoning in the production of scientific knowledge. In R. A. Duschl & R. J. Hamilton (Eds.), *Philosophy of science, cognitive psychology, and educational theory and practice* (p. 195-225). Albany, NY: State University of New York Press.

Norris, S. P., Falk, H., Federico-Agraso, M., Jiménez-Aleixandre, M. P., Phillips, L. M., Yarden, A. (2009a). Reading science texts – epistemology, inquiry, authenticity – A rejoinder to Jonathan Osborne. *Research in Science Education*, Published Online. doi:10.1007/s11165-008-9118-5

Norris, S. P., Macnab, Wonham, M., & de Vries, G. (2009b). West Nile virus: Using adapted primary literature in mathematical biology to teach scientific and mathematical reasoning in high school. *Research in Science Education,* Published Online. doi:10.1007/s11165-008-9112-y

Norris, S. & Phillips, L. (1994a). The relevance of a reader's knowledge within a perspectival view of reading, *Journal of Reading Behavior, 26,* 391-412.

Norris, S., & Phillips, L. (1994b). Interpreting pragmatic meaning when reading popular reports of Science. *Journal of Research in Science Teaching, 31,* 947-967.

Norris, S. P., & Phillips, L. M. (2001). How literacy in its fundamental sense is central to scientific literacy, paper presented at *the Annual Conference of the National Association for Research in Science Teaching,* St. Louis, MO, 26-28 March, 2001.

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

Norris, S. P., & Phillips, L. M. (2008). Reading as inquiry. In R. A. Duschl & R. E. Grandy (Eds.), *Teaching scientific inquiry: Recommendations for research and implementation* (pp. 233-262). Rotterdam, Netherlands: Sense.

Novak, J. D. (1988). Learning science and science of learning. *Studies in Science Education, 15*(1), 77-101.

Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn.* Cambridge, England: Cambridge University Press.

Nuffield Foundation (1966). *Nuffield chemistry: Introduction and guide*. Harlow, England: Longman/Penguin.

Office for Civil Rights (ED). (2012). *Gender equity in education: A data snapshot.* U.S. Department of Education. (ED 533168).

Ogborn, J. (1994). *Theoretical and empirical investigations of the nature of scientific and commonsense knowledge.* (Unpublished doctoral thesis). King's College, London.

Olson, D. R. (1994). *The world on paper*. Cambridge: Cambridge University Press.

O'Neal, C., Wright, M., Cook, C., Perorazio, T., and Purkiss, J. (2007). The Impact of teaching assistants on student retention in the sciences: Lessons for TA training. *Journal of College Science Teaching, 36*(5), 24-29.

Onwuegbuzie, A. J. (1995). Statistics test anxiety and female students. *Psychology of Women Quarterly, 19*, 413-418.

Onwuegbuzie, A. J., & Teddlie, C. (2003). A framework for analyzing data in mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioural research* (pp. 351-383). Thousand Oaks, CA: Sage.

Onwuegbuzie, A. J., & Johnson, R. B. (2006). The validity issue in mixed research. *Research In The Schools, 13*(1), 48-63.

Organisation for Economic Cooperation and Development [OECD] (2003). *PISA assessment framework – mathematics, reading, science and problem solving knowledge and skills.* Paris, France: Author.

Osborne, J. (1996a). Beyond constructivism. *Science Education, 80*(1), 53-82.

Osborne, J. (1996b). Untying the Gordian knot: diminishing the role of practical work. *Physics Education, 31*(5), 271-278.

Osborne, J. (2002). Science without literacy: A ship without a sail? *Cambridge Journal of Education, 32*(2), 203-218.

Osborne, J. (2009). The potential of adapted primary literature (APL) for learning: a response. *Research in Science Education, 39*(3), 397-403. doi:10.1007/s11165-008-9117-6

Osborne, J. (2010). Arguing to learn in science: The role of collaborative critical discourse. *Science, 328*, 463.

Osborne, J. (2011). Science teaching methods: a rationale for practices. *School Science Review*, *93*(343), 93-103.

Osborne, J. (2013). The 21st century challenge for science education: Assessing scientific reasoning. *Thinking Skills and Creativity, 10*, 265-279

Osborne, J., Driver, R. & Simon, S. (1998). Attitudes to science: Issues and concerns. *School Science Review, 79*(288), 27-33.

Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching, 41*(10), 994-1020.

Osborne, J., & Ratcliffe, M. (2002). Developing effective methods of assessing ideas and evidence. *School Science Review, 83*(305), 113-123.

Osborne, R. J., & Wittrock, M. C. (1983). Learning science: A generative process. *Science Education, 67,* 489-508.

Osborne, R. J., & Wittrock, M. C. (1985). The generative learning model and its implications for science education. *Studies in Science Education, 12,* 59-87.

Paivio, A. (1986). *Mental representations: A dual-coding approach*. New York, NY: Oxford University Press.

Paivio, A. (1990). *Mental representations: A dual-coding approach*. New York: Oxford University Press.

Pall, M. L. (2000). The value of scientific peer-reviewed literature in a general education science course. *The American Biology Teacher, 62*(4), 256-258.

Pardi, P. (2011, September 22). What is knowledge? *Philosophy News.* Retrieved from http://www.philosophynews.com/post/2011/09/22/What-is-Knowledge.aspx

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Peck, W. H. (2004). Teaching metastability in petrology using guided reading from the primary literature. *Journal of Geoscience Education, 52*(3), 284-288.

Penney, K., Norris, S. P., Phillips, L. M., & Clark, G. (2003). The anatomy of junior high school science textbooks: An analysis of textual characteristics and a comparison to media reports of science. *Canadian Journal of Science, Mathematics and Technology Education, 3*(4), 415-436.

Perkins, D. N., & Simmons, R. (1988). Patterns of misunderstanding: An integrative model for science, math and programming. *Review of Educational Research, 58*(3), 303-326.

Peshkin, A. (1993). The goodness of qualitative research. *Alberta Educational Researcher, 22*(2), 23-29.

Peterson, B. A., & Seligman, E. P. (1985). Depression symptoms and unprompted causal attributions: Content analysis. *Behavior Research and Therapy, 23*, 379-79-382.

Petri, C. (1999). *Traditionella och icke-traditionella studenter* [Traditional and non-traditional students; in Swedish]. StudS Arbetsrapport nr 8. Retrieved from www.hsv.se/verksamhet/statistik/studs

Pfundt, H., & Duit, R. (1994). *Bibliography: Students' alternative frameworks and science education*. Kiel, Germany: IPN.

Pfundt, H., & Duit, R. (1997). *Bibliography: Students' alternative frameworks and science education* (4th ed.). Kiel, Germany: University of Kiel.

Pfundt, H., & Duit, R. (2000). *Bibliography: Students' alternative frameworks and science education,* 5th edition. Kiel, Germany: University of Kiel.

Phillips, D. C. (Ed.) (1995). The good, the bad and the ugly: The many faces of constructivism. *Educational Researcher, 24*(7), 5-12.

Phillips, D. C. (2000). *Constructivism in education: Opinions and second opinions on controversial issues.* Chicago, IL: University of Chicago Press.

Phillips, D. C., & Burbules, N. C. (2000). *Positivism and educational research*. Oxford, England: Rowman & Littlefield Publishers, Inc.

Phillips, L. M., & Norris, S. P. (1999). Interpreting popular reports of science: What happens when the reader's world meets the world on paper? *International Journal of Science Education, 21*(3), 317-327.

Phillips, L. M., & Norris, S. P. (2009). Bridging the gap between the language of science and the language of school science through the use of adapted primary literature. *Research in Science Education*, *39*(3), 313-319. doi:10.1007/s11165-008-9111-z

Phillips, L. M., Yarden, A., Falk, H., Norris, S. P., Jimenez-Aleixandre, M. P., & Ford, D. J. (2008). *Reading scientific texts: Adapting primary literature for promoting scientific* *literacy.* (Symposium presented at the NARST Annual International Conference: Impact of Science Education Research on Public Policy). Baltimore, MD.

Phinney, J. (1989). Stages of ethnic identity development in minority group adolescents. *Journal of Early Adolescence, 9*, 34-49.

Phinney, J., Ferguson, D., & Tate, J. (1997). Intergroup attitudes among ethnic minority adolescents: A causal model. A causal model. *Child Development, 68*, 955-969.

Phinney, J. S., Jacoby, B., & Silva, C. (2007). Positive intergroup attitudes: The role of ethnic identity. *International Journal of Behavioral Development, 31*(5), 478-490.

Piaget, J. (1950). *The psychology of intelligence*. London, England: Routledge & Kegan Paul.

Piaget, J. (1966). *The psychology of intelligence.* Totawa, NJ: Littlefield, Adams.

Piaget, J. (1978). *The development of thought* (A. Rosin, Trans.). Oxford, England: Basil Blackwell.

Piaget, J. (1985). *The equilibration of cognitive structures: The central problem of intellectual development* (T. Brown & K. J. Thampy, Trans.). Chicago, IL: University of Chicago Press. (Original work published 1975).

Polanyi, M. (1958). *Personal knowledge*. London, England: Routledge & Kegan Paul.

Pope, M. (1985). *Constructivist goggles: Implications for process in teaching and learning.* Paper presented at the annual conference of the British Educational Research Association, University of Sheffield.

Popper, K. [1945] (1985). Against the sociology of knowledge. In B. Magee (Ed.), *Popper selections*. Princeton, NJ: Princeton University Press.

Popper, K. R. (1959). *The logic of scientific discovery*. London: Hutchinson & Co Ltd.

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*, 211-227. doi: 10.1002/sce.3730660207

Posner, M. I., & McLeod, P. (1982). Information processing models, in search of elementary operations. *Annual Review of Psychology, 33*, 477-514.

Project Kaleidoscope (1991). *What works: Building natural science communities*. Washington, D. C.: Project Kaleidoscope.

Pryor, J. (2001). Highlights of recent epistemology. *British Journal for the Philosophy of Science, 52*, 95-124.

Puntambekar, S. (2009). *Scaffolding*. Retrieved January 9, 2016 from: http://www. education.com/reference/articles/scaffolding/

Ratcliffe, M. (1999). Evaluation of abilities in interpreting media reports of scientific research. *International Journal of Science Education, 21*(10), 1085-1099.

Ravetz, J. R. (1971). *Scientific knowledge and its social problems.* Oxford, England: Oxford University Press.

Read, J. R. (2004). *Children's misconceptions and conceptual change in science education.* Retrieved from: http://acell.chem.usyd.edu.au/Concelptual-Change.cfm

Reichardt, C. S., & Rallis, S. F. (1994). Qualitative and quantitative inquiries are not incompatible: A call for a new partnership. In C. S. Reichardt & S. F. Rallis (Eds.), *The qualitative-quantitative debate: New perspectives* (pp. 85-92). San Francisco, CA: Jossey-Bass.

Reif, F. (1987). Instructional design, cognition, and technology: Applications to the teaching of science concepts. *Journal of Research in Science Teaching, 24,* 309-324.

Reinharz, S. (1992). *Feminist methods in social research*. Oxford, England: Oxford University Press.

Reiss, M. (1998). Science for all. In M. Ratcliffe (Ed.), *ASE Guide to Secondary Science Education*. Hatfield, England: ASE.

Reteguiz, J-A. (2006). Relationship between anxiety and standardized patient test performance in the medicine clerkship. *Journal of General Internal Medicine, 21*(5), 415-418. doi: 10.111/j.1525-1497.2006.00419.x

Riffe, D., & Freitag, A. (1997). A content analysis of content analyses: Twenty five years of Journalism Quarterly. *Journalism and Mass Communication Quarterly, 74*, 635-644.

Roberts, R. (2001). Procedural understanding in biology: The 'thinking behind the doing.' *Journal of Biological Education, 35*(3), 113-117.

Roberts, R., & Gott, R. (1999). Procedural understanding: Its place in the biology curriculum. *School Science Review, 81*(294), 19-25.

Roberts, R., & Gott, R. (2000). Procedural understanding in biology: How is it characterized in texts? *School Science Review, 82(298)*, 83-91.

Roberts, R., & Gott, R. (2004). A written test for procedural understanding: A way forward for assessment in the UK science curriculum? *Research in Science and Technological Education, 22*(1), 5-15.

Roberts, R., Gott, R., & Glaesser, J. (2010). Students' approaches to open-ended science investigation: the importance of substantive and procedural understanding. *Research Papers in Education, 25*(4), 377- 407.

Rodrigues, S., Jindal-Snape, D., & Snape, J. B. (2011). Factors that influence student pursuit of science careers; the role of gender, ethnicity, family and friends. *Science Education International, 22*(4), 266-273.

Roecker, L. (2007). Introducing students to the scientific literature. *Journal of Chemical Education*, *84*(8), 1380-1384.

Rosser, S. V. (2004). *The science glass ceiling: Academic women scientists and the struggle to succeed.* Oxford, England: Taylor & Francis.

Roth, W-M. (2001). "Enculturation": Acquisition of cultural blind spots and epistemological prejudices. *British Educational Research Journal, 27*(1), 5-27.

Roth, W-M., & Roychoudhury, A. (1993). The development of science process skills in authentic contexts. *Journal of Research in Science Teaching, 30*(2), 127-152.

Ruble, D., Alvarez, J., Bachman, M., Cameron, J., Fuligni, A., & Garcia Coll, C. (2004). The development of a sense of "we": The emergence and implications of children's collective identity. In M. Bennett & F. Sani (Eds.), *The development of the social self* (pp. 29-76). New York, NY: Psychology Press.

Ryder, J., Leach, J., & Driver, R. (1999). Undergraduate science students' images of science. *Journal of Research in Science Teaching, 6*(2), 201-220.

Ryder, J. & Leach, J. (2000). Interpreting experimental data: The view of upper secondary school and university science students. *International Journal of Science Education, 22*(10), 1069-1084.

Samarapungavan, A. (1992). Children's judgments in theory choice tasks: Scientific rationality in childhood. *Cognition, 45*(1), 1-32.

Sandelowski, M. (2001). Real qualitative researchers don't count: The use of numbers in qualitative research. *Research in Nursing & Health, 24*, 230-240.

Sandelowski, M. (2003). Tables or tableux? The challenges of writing and reading mixed methods studies. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 321-350). Thousand Oaks, CA: Sage.

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

Sandoval, W. A., & Millwood, K. A. (2005). The quality of students' evidence in written scientific explanations. *Cognition and Instruction, 23*(1), 23-55.

Schauble, L., & Kuhn, D. (1989). *Applying scientific thinking skills in self-directed exploration of a microworld.* Manuscript submitted for publication. (Available from L. Schauble, LRDC, University of Pittsburgh, Pittsburgh, PA 15260).

Schellings, G.L.M., van Hout-Wolters, B.H.A.M., & Vermunt, J.D. (1996). Selection of main points in instructional texts: Influences of task demands. *Journal of Literacy Research, 28,* 355-378.

Schmidt, P. (2002). States push public universities to commercialize research. *Chronicle of Higher Education, 48*:A26.

Schwab, J. J. (1962). The teaching of science as inquiry. In J. J. Schwab & P. F. Brandwein (Eds.). *The teaching of science*. Cambridge: Harvard University Press.

Schwandt, T. A. (1998). Constructivist, interpretivist approaches to human inquiry. In N. K. Denzin & Y. S. Lincoln (Eds.), *The landscape of qualitative research: Theories and issues* (pp. 221-259). London, England: Sage.

Science Council, United Kingdom. (2010). *What is science?* Retrieved from http://www.sciencecouncil.org/definition

Scriven, M. (1971). Objectivity and subjectivity in educational research. In L. G. Thomas (Ed.), *Philosophical redirection of educational research* (71st Yearbook of the National Society for the Study of Education, Part 1). Chicago, IL: University of Chicago Press.

Scott, D. (2005). Retention, completion and progression in tertiary education in New Zealand. *Journal of Higher Education Policy and Management, 27*(1), 3-17.

Scruggs, M. M. (1988). What research says about textbooks. *Science and Children, 25*(4), 24-25.

Sengupta, P., Krinks, K. D., & Clark, D. B. (2015). Learning to deflect: Conceptual change in physics during digital game play. *Journal of the Learning Sciences, 24*, 638-674.

Séré, M.-G., Journeaux, R., & Larcher, C. (1993). Learning the statistical analysis of measurement error. *International Journal of Science Education, 15*, 427-438.

Seymour, E., & Hewitt, N. M. (1997). *Talking about leaving: Why undergraduates leave the sciences*. Boulder, CO: Westview Press.

Seymour, E., Hunter, A. B., Laursen, S. L., & Deantoni, T. (2004). *Science Education, 88*(4), 493-534,

Shanahan, M. C., Delos Santos, J. S., & Morrow, R. (2009). Hybrid adapted primary literature: A strategy to support elementary students in reading about scientific inquiry. *Alberta Science Education Journal, 40*(1), 20-26.

Shuell, T. J. (1996). Teaching and learning in a classroom context. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 726-764). New York, NY: Mac-millan.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*, 4-14.

Sieber, S. D. (1973). The integration of fieldwork and survey methods. *American Journal of Sociology, 73*, 1335-1359.

Sjøbeg, S., & Schreiner, C. (2005). *Young people and science attitudes, values and priorities: evidence from the ROSE project*. Keynote presentation at the European Union Science and Society Forum, Brussels, Belgium.

Slavin, R. E. (1988). *Educational psychology: Theory into practice.* Englewood Cliffs, NJ: Prentice Hall.

Smith, C. P., (Ed.) in association with Atkinson, J. W., McClelland, D. C., & Veroff, J. (1992). *Motivation and personality: Handbook of thematic content analysis*. Cambridge, MA: Cambridge University Press.

Smith, G. R. (2001). Guided literature exploration: Introducing students to the primary literature. *Journal of College Science Teaching, 30*, 465-469.

Smith, J. K., & Heshusius, L. (1986). Closing down the conversation: The end of the quantitative-qualitative debate among educational researchers. *Educational Researcher, 15*, 4-12.

Snyder, S. A., & Dillow, T. D. (2009). *Digest of Education Statistics 2009,* NCES 2010-013, Washington, DC: NCES.

Solomon, J. (1994). The rise and fall of constructivism. *Studies in Science Education, 23*, 1-19.

Stake, R. E., & Easley, J. A. (1978). *Case studies in science education.* Urbana: University of Illinois Centrefor Instructional Research and Curriculum Evaluation.

Stanley, L., & Wise, S. (1983). *Breaking out: Feminist consciousness and feminist research.* London, England: Routledge & Kegan Paul.

Stenhouse, L. (1975). *An Introduction to Curriculum Research and Development*. London, England: Heinemann

Sternberg, R. (1985). *Beyond IQ:A triarchic theory of human intelligence*. Cambridge, England: Cambridge University Press.

Strike, K. A., Posner, G. J. (1992). A revisionist theory of conceptual change. In R. A. Duschl & R. J. Hamilton, (Eds.), *Philosophy of science, cognitive psychology, and educational theory and practice* (pp. 147-146). Albany, NY: SUNY Press.

Stuart, V. (2000). Math curse or math anxiety? *Teaching Children Mathematics, 6*(5), 330-335.

Suppe, F. (1998). The structure of a scientific paper. *Philosophy of Science, 65,* 381-405.

Sylvester, R. (1995). *A celebration of neurons: An educator's guide to the human brain.* Alexandria, VA: Association for supervision and curriculum development (ASCD).

Tabak, I., & Baumgartner, E. (2004). The teacher as partner: Exploring participant structures, symmetry, and identity work in scaffolding. *Cognition and Instruction, 22*(4), 393-429.

Tai, R. H., Liu, C. Q., Maltese, A. V., & Fan, X. (2006). Planning early for careers in science. *Science, 312*, 1143.

Tajfel, H., & Turner, J. (1986). The social identity theory of intergroup behaviour. In S. Worchel & W. Austin (Eds.), *Psychology of intergroup relations* (pp. 7-24). Chicago, IL: Nelson-Hall.

Tashakkori, A., & Creswell, J. W. (2007). The new era of mixed methods. *Journal of Mixed Methods Research, 1*, 3-7.

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining the qualitative and quantitative approaches.* Thousand Oaks, CA: Sage.

Taylor, J. A., & Dana, T. M. (2003). Secondary school physics teachers' conceptions of scientific evidence: An exploratory study. *Journal of Research in Science Teaching, 40,* 721-736.

Taylor, N., & Coll, R. (1997). The use of analogy in the teaching of solubility to preservice primary teachers. *Australian Science Teaching Journal, 43*(4), 58-64.

Teddlie, C., & Tashakkori, A. (2003). Major issues and controversies in the use of mixed methods in the social and behavioral sciences. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 3-50). Thousand Oaks, CA: Sage.

Teddlie, C., & Tashakkori, A. (2009*). Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences.* Thousand Oaks, CA: Sage.

Tezcan, H., & Yilmaz, Ü. (2003). The effect of conceptual computer animation and traditional instruction in teaching chemistry on student achievement. *Journal of the Faculty of Education, 14*(2), 18-32.

The President's Council of Advisors on Science and Technology (2011). K-12 science, technology, engineering and math (STEM) education for America's future. *Tech Directions, 70(6)*, 33-34.

Thomas, K. (2008). Evidence. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*, Retrieved from http://plato.stanford.edu/archives/fall2008/entries/evidence/

Tobias , S. (1992). Women in science: Women and science. *Journal of College Science Teaching, 21*(5), 276-278.

Tobin, K. (Ed.). (1993). *The practice of constructivism in science education.* Hillsdale, NJ: Lawrence Erlbaum.

Tobin, K., & Tippins, D. (1993). Constructivism as a referent for teaching and learning. In K. Tobin (Ed.). *The practice of constructivism in science education.* Washington, DC: AAAS Press, pp. 3-21.

Toulmin, S. E. (1958). *The uses of argument*. Cambridge, England: Cambridge University Press.

Truth and Reconciliation Commission of Canada, (2015). *TRC Final Report.* Retrieved from http://www.trc.ca/websites/trcinstitution/index.php?p=3

Tucker, S. A., Hanuscin, D. L., & Bearnes, C. J. (2008). Igniting girls' interest in science. *Science 319,* 1621-1622.

Tuhiwai Smith, L. (1999). *Decolonizing methodologies: Research and Indigenous peoples*. London, England: Zed Books.

Tytler, R. (2002). Teaching for understanding in science: Constructivist/conceptual change teaching approaches. *Australian Science Teaching Journal, 48*(4), 30-35.

Udo, M. K., Ramsey, G. P., & Mallow, J. V. (2004). Science anxiety and gender in students taking general education science courses. *Journal of Science Education and Technology, 13*(4), 435-446.

University of Alberta, (n.d.). *Science 100*. Retrieved from https://science.ualberta.ca/programs/undergraduate/science-100

U.S. Government Accountability Office. (2004). *Report to congressional requesters: Gender issues: Women's participation in the Sciences has increased, but agencies need to do more to ensure compliance with Title IX,* GAO-04-639. Washington, DC: U.S. Government Accountability Office.

Vanderwolf, C. H., Cook, M., Coutts, R. T., & Cropp, D. (2005). *Teaching science in the 21st century: An examination of Canadian science curricula from kindergarten to Grade 12*. A report prepared for the Society for Quality Education, Waterloo, Ontario.

Van Driel, J. H. (2002). Students' corpuscular conceptions in the context of chemical equilibrium and chemical kinetics. *Chemical Education Research Practice in Europe, 3*(2), 201-213.

van Lier, L. (1996). *Interaction in the language curriculum*. New York, NY: Longman.

van Manen, M. (2007). *Researching lived experience: Human science for an action sensitive pedagogy*. Ontario, Canada: The Althouse Press.

van Peer, W., & Chatman, S. (Eds.). (2001). *New perspectives on narrative perspective*. Albany, NY: State University of New York Press.

Vartalla, T. (1999). Remarks on the communicative functions of Hedging in popular scientific and specialist research articles on medicine. *English for Specific Purposes, 18*(2), 177-200.

Viennot, L. (1979). Spontaneous reasoning in elementary dynamics. *European Journal of Science Education, 1*, 205-221.

Vinge, V. (1993). *The technological singularity: How to survive in the post-human era*. Retrieved from: *www-rohan.sdsu.edu/faculty/vinge/misc/singularity.html*

von Glasersfeld, E. (1991). Cognition, construction of knowledge, and teaching. In M. R. Mathews (Ed.), *History, philosophy and science teaching* (pp. 117-132). New York, NY: Teachers College Press..

Vosniadou, S. (2007). Conceptual change and education, *Human Development, 50*, 47-54. doi: 10.1159/000097684

Vygotsky, L. S. (1962). *Thought and language*. Cambridge, MA: MIT Press.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological process.* Cambridge, MA: Harvard University Press.

Walker, I. (2007). *Null hypothesis testing and effect sizes. Statistics for Psychology: Making sense of our world through analyses.* Retrieved from staff.bath.ac.uk/pssiw/stats2/page2/page14/page14.html

Watson, R., Swain, J., & McRobbie, C. (2004). Students' discussions in practical scientific enquiries. *International Journal of Science Education, 26*(1), 25-46.

Weaver, C. A., & Kintsch, W. (1991). Expository text. In R. Barr, M. I. Kamil, P. Mosenthal, & P. D. Pearson (Eds.). *Handbook of reading research* (Vol. 2, pp. 230-245). White Plains, NY: Longman.

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). *Unobtrusive measures: Nonreactive research in the social sciences.* Chicago IL: Rand McNally.

Weber, S. J. (1986). The nature of interviewing. *Phenomenology and Pedagogy, 4*(2), 65-72.

Weisgram, E. S., & Bigler, R. S. (2007). Effects of learning about gender discrimination on adolescent girls' attitudes toward and interest in science. *Psychology of Women Quarterly, 31,* 262-269.

Weiss, I. R., Banilower, E. R., McMahon, K. C., & Smith, P. S. (2001). *Report of the 2000 National Survey of Science and Mathematics Education.* Chapel Hill, NC: Horizon Research, Inc., 2001.

Wellington, J. (1991). Newspaper science, school science: friends or enemies? *International Journal of Science Education, 13*(4), 363-372.

Wenger, D. M. & Loppnow, G. R. (2004). Applications: Broadband photochemistry for the origins of life. *The Spectrum, 17*(4), 26-29.

Wertsch, J. V. (1991). A sociocultural approach to socially shared cognition. In L. B. Resnick, J. M. Levine, & S. D. Teasly (Eds.). *Perspectives on socially shared cognition* (pp. 85-100). Washington, DC: American Psychological Association.

White, R. (1988). *Learning science*. Oxford, England: Basil Blackwell.

Widodo, A., Duit, R., & Müller, C. (2002). *Constructivist views of teaching and learning in practice: Teachers' views and classroom behaviour.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, New Orleans.

Wiens, D. J., Depping, D. J., Wallerich, S. R., Van Laar, E. S., & Juhl, A. L. (2003). Gender matters: Factors influencing biology interest among undergraduate majors. *Journal of College Science Teaching, 33*(1), 32-36.

Wilson, E. (2002). Learning concepts. In P. Warwick & R. S. Linfield (Eds.), *Science 3-13. The Past, the Present and Possible Futures* (pp. 37-49). New York, NY: Routledge, Taylor & Francis Group.

Wilson, S. (2001). Editorial: Self-as-relationship in Indigenous research. *Canadian Journal of Native Education, 25*(2), 91-92.

Wilson, K. & Korn, J. H. (2007). Attention during lectures: Beyond ten minutes. *Teaching of Psychology, 34(*2), 85-89.

Wimmer, R., Legare, L., Arcand, Y., & Cottrell, M. (2009). Experiences of beginning Aboriginal teachers in band-controlled schools. *Canadian Journal of Education, 32* (4), 817-849.

Wimmer, R. (2013). *The four Rs revisited, again: Research and reflections on Aboriginal teacher education and educational leadership in Higher Education*. BELMAS (British Educational Leadership, Management and Administration Society) Annual conference, July, 2013. Edinburgh, Scotland, UK.

Wiser, M., & Amin, T. G. (2002). Computer-based interactions for conceptual change in science. In M. Limon & L. Mason (Eds.), *Reconsidering conceptual change: Issues in theory and practice* (pp. 357-387). Dordrecht: Kluwer

Wonham, M., de-Camino-Beck, T., & Lewis, M. (2004). An epidemiological model for West Nile virus: Invasion analysis and control applications. *Proceedings of the Royal Society of London B, 271,* 501-507.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 17*(2), 89-100.

Wood, D., & Wood, H. (1996). Vygotsky, tutoring and learning. *Oxford Review of Education, 22*(1), 5-16.

Worchel, S. (1999). *Written in blood: Ethnic identity and the struggle for human harmony*. New York, NY: Worth.

Wu, H. K., & Krajcik, J. (2006). Inscriptional practices in two inquiry-based classrooms: A case study of seventh graders' use of data tables and graphs. *Journal of Research in Science Teaching, 43*(1), 63-95.

Yager, R. E. (1983). The importance of terminology in teaching K-12 science. *Journal of Research in Science Teaching, 20,* 577-588.

Yarasi, S., McConachie, C., & Loppnow, G. R. (2005). Molecular beacon probes of photodamage in thymine and uracil oligonucleotides. *Photochemistry and Photobiology, 81*, 467-473.

Yarden, A. (2009). Reading scientific texts: Adapting primary literature for promoting scientific literacy. *Research in Science Education, 39,* 307-311.

Yarden, A., Brill, G., & Falk, H. (2001). Primary literature as a basis for a high-school biology curriculum. *Journal of Biological Education, 35*(4), 190-195.

Yarden, A., Falk, H., Federico-Agraso, M., Jiménez-Aleixandre, M. P., Norris, S. P., Phillips, L. M. (2009). Supporting teaching and learning using authentic scientific texts: A rejoinder to Danielle J. Ford. *Research in Science Education, 39,* 391-395.

Yarden, A., Norris, S. P., & Phillips, L. M. (2015). *Adapted Primary Literature: The use of authentic scientific texts in secondary schools.* Dordrecht, The Netherlands:Springer.

Yolles, M. (2006). *Organizations as complex systems: An introduction to knowledge cybernetics.* New York, USA: Information Age Publishing.

Yore, L. D., & Treagust, D. F. (2006). Current realities and future possibilities: Language and science literacy – empowering research and informing instruction. *International Journal of Science Education, 28*(2-3), 291-314.

Yuenyong, C., & Narjaikaew, P. (2009). Scientific literacy and Thailand science education. *International Journal of Environmental Science Education, 4*(3), 335-349.

Zalta, E. N. (Ed.). (2006). *The Stanford encyclopedia of hilosophy*. Retrieved from http://plato.stanford.edu/entries/evidence

Zhou, G. (2010). Conceptual change in science: A process of argumentation. *Eurasia Journal of Mathematics, Science & Technology Education, 6*(2), 101-110.

Zhou, G., Brouwer, W., & Nocente, N. (2005). Enhancing conceptual learning through computer-based applets: The effectiveness and implications. *Journal of Interactive Learning Research, 16*(1), 31-49.

Zull, J. E. (2004). The art of changing the brain. *Educational Leadership, 62*(1), 68-72.

Zumdahl, S. A. (1996). Mission impossible? Improving retention of science majors among minorities and women. *Journal of Chemical Education, 73*(11), A266-A267.

Appendix A

Adapted Primary Literature

ADAPTED PRIMARY LITERATURE FROM

Yarasi, S., McConachie, C., and Loppnow, G.R. (2005). Molecular beacon probes of photodamage in thymine and uracil oligonucleotides. *Photochemistry and Photobiology, 81*, 467-473.

Introduction

When the organic bases in nucleic acids contained within cells are exposed to ultraviolet (UV) light, unwanted biological effects occur in them. Some of these harmful effects lead to death, mutations and even cancer. When UV light falls on these cells the nucleic acids are damaged forming photoproducts, which are primarily cyclobutane pyrimidine dimers (CPD), photohydrates and pyrimidine-pyrimidone (6-4) products. Of these three photoproducts, CPDs are the most abundant. When these photoproducts are formed, lesions appear in the DNA. These lesions inhibit the enzymes responsible from carrying out the replication and transcription of DNA. The DNA lesions are also believed to be carcinogenic because they could lead to miscoding during replication. The effect of the corresponding lesions in the functioning of RNA is unknown. There has been a renewed interest in the molecular basis of the carcinogenic effects of UV light because of the recent awareness of the depletion of the stratospheric ozone layer, which is used to effectively screen out most of the biologically damaging UV light in

- Organic bases - guanine, cytosine, adenine and thymine – when bound to sugar and phosphate groups are referred to as nucleotides. When these single units called nucleotides are joined together they produce the complex molecule (polymer) called DNA (Deoxyribonucleic acid).
- Cyclobutane pyrimidine dimers (CPD) are made of two thymine molecules, connected together in such a way as to form a cyclobutane ring between them.
- Photohydrates are the addition products formed between a pyrimidine and water, and are when the UV-light is absorbed by the pyrimidine bases in DNA and RNA. Refer to Figure 2 in the Appendix for the structure/formation of photohydrates.
- Pyrimidine-pyrimidone(6-4) products – are formed when UV-light falls on the DNA. These photoproducts are formed similarly as CPDs, but are slightly different structurally.
- Lesion – a wound or injury.
- Miscoding – insertion of wrong base when DNA undergoing replicatio
- UV light - . The electromagnetic spectrum of ultraviolet light can be subdivided into UVA (long wave), UVB (medium wave) and UVC

the 200-400 nm region.

In order to understand the killing and mutation- inducing effects of photoproducts generated by UV light, it is essential to have methods that can measure DNA damage and repair in a sensitive and precise manner. Various techniques, such as gel electrophoresis, HPLC and mass spectrometry - to name a few - have been developed to measure the damage caused to DNA by chemical agents, UV light and ionizing radiation. All of these methods have their own advantages and disadvantages. However, all these techniques destroy human cells, and therefore cannot be adapted for *in vivo* or *in situ* use.

In recent years, a number of fluorescence methods have been proposed for probing DNA damage. These fluorescence methods are more sensitive and can be used both *in vivo* and *in situ.*

Molecular beacons (MB) are a relatively new class of nucleic acid probes which function effectively because:

1) They hybridize with complementary nucleic acid bases in their specific sequence
2) They fluoresce when energy is released from them.

(short wave) which show a corresponding decrease in the wavelength range in nanometers, and an increase in the energy per photon. The Sun emits ultraviolet radiation in the UVA. UVB, and UVC bands. However, because of the absorption in the atmosphere's ozone layer, 98.7% of the ultraviolet light that reaches the Earth's surface is UVA.

- 200-400 nm region – 1nm= 10^{-9}m
- Mutations – are changes to the nucleotide sequence of the genetic material of an organism.
- Gel electrophoresis – is a technique used for the separation of DNA and RNA, or protein molecules using an electric current applied to a gel matrix. The term "gel" refers to the matrix, which is like jello, used to contain and then separate the target molecules.
- HPLC – High-performance liquid chromatography is a form of column used to purify, separate, identify and quantify individual compounds from mixtures of compounds.
- Mass Spectrometry – is a technique that measures the mass-to-charge ratio of charged particles.
- *in vivo* – experimentation done in or on the living tissue of a whole, living organism. Animal testing and clinical trials are forms of *in vivo* research.
- *in situ* - means to examine the phenomenon exactly in the place where it occurs, without moving it to some special medium. For example, when a cell is examined within a whole organ, which is kept intact.
- Probing – studying, looking at with the purpose of understanding/detecting
- Molecular beacons – are single-stranded oligonucleotides having a stem-and-loop structure as shown in Figure 1 and in Table 1 of the Appendix. The loop contains a nucleotide sequence that is

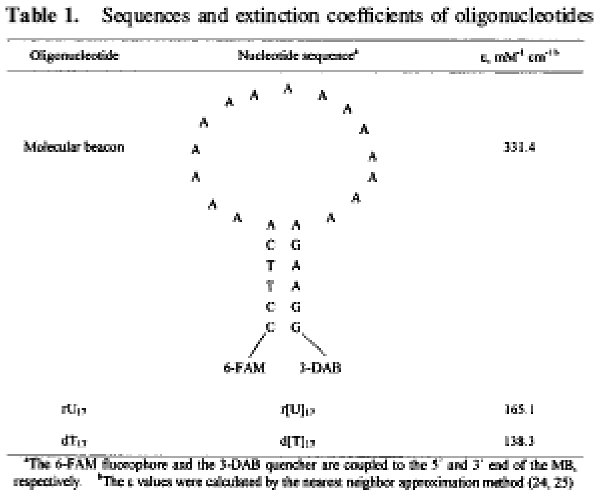


Table 1. Sequences and extinction coefficients of oligonucleotides

Oligonucleotide	Nucleotide sequence[a]	ε, mM^{-1} cm^{-1} [b]
Molecular beacon	6-FAM–CCTTCA AAAAAAAAAAAAAAA AGAAGG–3-DAB	331.4
rU$_{17}$	r[U]$_{17}$	165.1
dT$_{17}$	d[T]$_{17}$	138.3

[a] The 6-FAM fluorophore and the 3-DAB quencher are coupled to the 5' and 3' end of the MB, respectively. [b] The ε values were calculated by the nearest neighbor approximation method (24, 25)

MB are oligonucleotides which have a stem-loop (hairpin) structure as shown in Table 1, with a fluorophore and a nonfluorescent quencher attached to the 5' and 3' ends respectively. Alone, this probe has a hairpin conformation. The fluorophore and quencher then, are close to each other, they share energy or electrons and the fluorescence is quenched. However, when a complementary target sequence appears, the probe binds to the target sequence. This binding causes the fluorophore and the quencher to separate. The fluorophore then absorbs light and reradiates the energy as visible light, resulting in an increased fluorescence. This increase in fluorescence is a signal which can be easily detected and measured. MB are, therefore, highly selective and sensitive probes which have three advantages:

complementary to a target sequence. A fluorophore is covalently linked to the 5'end of one arm and a quencher is covalently linked to the 3'end of the other arm. Molecular beacons do not fluoresce when they are free in solution. However, when they hybridize to a nucleic acid strand (a target), they undergo a conformational change that enables them to fluoresce brightly.

- Hybridize – to form base pairs between complementary regions of two strands of DNA that were not originally paired.
- Fluoresce – absorb radiation and reradiate the energy as light of lower energy.
- Oligonucleotides – a short polymer of two to twenty nucleotides.
- Stem-loop (hairpin structure) – Shown as Figure 1 in Appendix.
- Fluorophore – a molecule. or a part of a molecule, which fluoresces.
- Non-fluorescent quencher – the part of the MB which accepts electrons or energy.
- 5' and 3' ends – The two ends of a nucleotide chain.
- Target – The nucleotide sequence to which the loop of the MB binds.
- Fluorophore – is a component/part of a molecule which causes it to fluoresce. It is a part in a molecule which will absorb energy of a specific wavelength and re-emit it at a lower energy, but at a higher (but equally specific) wavelength.

- They can recognize target sequences that differ by a single mismatched base pair.
- The measurements are rapid and easy
- They have the potential to be applied *in vivo.*

MB have never before been used to detect nucleic acid damage induced by UV light. This study reports a new fluorescent probe, based on MB, used for detecting UV-induced photoproducts formed in DNA and RNA. This new method is based on the principle that when nuclei acids are damaged there is a disruption in their base pairing. If this disruption is large enough it can lead to a loss in fluorescence. MB can be used to detect nucleic acid damage only if the disruption in base pairing is big enough to produce a measurable loss in fluorescence.

MBs are very specific probes for detecting base-pair mismatches, and could potentially be used to detect photoproduct formation in oligonucleotides, especially poly-rU. The sensitivity they show indicates that MB have the ability to detect all types of nucleic acid (i.e. DNA and RNA) damage, modification and repair.

Materials and Methods

Materials.

The structure of the MB that was used in this study is

- Molar extinction coefficient or Extinction coefficient – is the measure of how strongly a substance shown in Table 1. The fluorophore, 6-FAM, and the quencher, 3-DAB are coupled to the 5' and 3' ends of the MB respectively. The single-strand target oligonucleotides that were used in this study are rU_{17} and dT_{17}. The extinction coefficients of the MB as well as of the two oligonucleotides are listed in Table 1. Nanopure water was used to make up all the solutions.

The oligonucleotide samples were separated into aliquots and frozen at -20^0C until needed. The concentrations of the oligonucleotides were determined by using the molar extinction coefficient calculated by utilizing the empirical formula given below. This formula takes into account the nearest neighbor interactions in short oligonucleotides.

The formula is:

$$\varepsilon_{260} = \sum_{1\rightarrow n-1}(\varepsilon_{\text{nearest neighbor}}) - \sum_{2\rightarrow n-1}(\varepsilon_{\text{individual}})$$

where,

ε_{260} = the molar extinction coefficient at 260nm of the oligonucleotide with a length of *n* bases

$\varepsilon_{\text{nearest neighbor}}$ = molar extinction coefficients of a pair of adjacent nucleobases

$\varepsilon_{\text{individual}}$ = molar extinction coefficients of individual bases

absorbs light at a particular wavelength, and is usually represented by the unit $M^{-1}cm^{-1}$. It is symbolized by ε in the Beer-Lambert law which states:

$$A = \varepsilon cl \quad \text{where,}$$

A= the absorbance of the sample
l = the pathlength (usually in cms)
c = the concentration of the absorbing species in the material (usually in moles per liter).

- Nanopure water - is even purer than deionized water in that organic and inorganic compounds have been removed from it by filtering through several filters.
- rU_{17} – a short polymer of 17 uracil ribonucleotides.
- dT_{17} – a short polymer of 17 thymine deoxyribonucleotides
- Aliquot – a small volume of sample which is analyzed and is assumed to be representative of the entire sample.
- Nearest neighbor interactions – the interactions between neighbouring sites which are closest together.

$\sum$ = the sums are taken over all possible sets of neighboring bases in the oligonucleotides

n = length of the oligonucleotide

Irradiations.

UV light obtained from a 1000 W Xenon lamp was passed through a monochromator. The resulting monochromatic, UV light was used to shine simultaneously on two 1 cm pathlength cuvettes at 280 nm. The cuvettes contained 3mL of 6.67 µM solutions of dT_{17} and rU_{17} which had been purged with nitrogen. The entrance and exit slit widths of the monochromator were adjusted to be 4.8nm. The power at the exit slit of the monochromator was approximately 5 to7 mW. Ozone generated from the lamp was ventilated away from the samples. The dT_{17} and rU_{17} solutions were stirred constantly during the period in which they were irradiated with UV light. The UV lamp was on throughout the experiment. The absorption spectra were recorded at intervals throughout the irradiation period. Control solutions were treated identically except that they were not exposed to the UV light source. Absorption spectra were recorded on spectrophotometers. Two types of spectrophotometers were used:

- A Cary 400 scan UV-Visible spectrophotometer

- Monochromator – a spectroscope with a slit that can be moved across the spectrum for viewing individual spectral bands. It is, therefore, a device for isolating a narrow portion of a spectrum.
- Pathlength – the distance that light (UV/VIS) travels through a sample in an analytical cell. In spectrophotometry the pathlength is measured in centimeters.
- Cuvette – a small, transparent, often a rectangular cubic laboratory vessel.
- Purged – bubbled into the sample.
- Irradiate – to expose to light
- Absorption spectrum – A plot of the amount of absorbed light by a medium as a function of wavelength or frequency.

OR

- A Hewlett-Packard 8452A diode array spectrophotometer.

Broadband irradiation was carried out using similar conditions and procedures except for the following details. A 200 W broadband mercury lamp was used to irradiate the samples. The mercury lamp was from Hanovia (Union, New Jersey). Across the UV and visible regions, a typical dose of light was approximately 2×10^{-2} W mm^{-2}, and roughly 10% of that dose was at wavelengths below 400nm. A quartz water jacket surrounded the lamp, to filter out the infra red radiation, and to prevent heating of the samples.

Fluorescence measurements.

The fluorescence spectra were measured using a spectrofluorimeter (from Shimadzu, Columbia, MD). Fluorescence emission spectra were recorded in the wavelength range between 490 and 700 nm. The excitation occurred at 480nm and bandwidths of 3 nm were used for both the excitation and the emission. A Suprasil quartz fluorescence cuvette, with a pathlength of 10mm and a volume of 300 µl, was used for these measurements. Throughout the irradiation, at regular intervals, an aliquot of the irradiated solution was mixed

- Spectrophotometer – an instrument used to determine the intensity of various wavelengths in a spectrum of light. It is a photometer measuring the relative intensities of the light in different parts of a spectrum.
- Diode array spectrophotometer – is a current model economical system for UV-visible spectroscopy based on a photodiode array detector.
- Photodiode array detector (PDA) – is a linear array of discrete photodiodes on an integrated circuit (IC) chip. In a spectrophotometer, it is placed at its image plane, in order to allow a range of wavelengths to be detected simultaneously.
- Fluorescence spectra – Fluorescence spectroscopy analyzes light emitted from a sample. It involves using a beam of light, usually ultraviolet light, that excites the electrons in molecules of certain compounds and causes them to emit light of a lower energy, typically, but not necessarily, visible light. The species being examined will have a ground electronic state (a low energy state) of interest, and an excited electronic state of higher energy
- Spectrofluorimeter – is an instrument in which the spectrum of emitted fluorescent light is measured.
- Excitation – a process in which a molecule, atom, nucleus or particle absorbs energy.
- Bandwidth – the numerical difference between the upper and lower frequencies of a band of electromagnetic radiation.

with an aliquot of 5 μM molecular beacon solution. This mixture was then incubated in a water bath for 3 hours at 27-29^{0}C for rU$_{17}$, and for 1.5 hours at 30-32^{0}C for dT$_{17}$. The fluorescence spectra were recorded at room temperature using a 200 μL sample of the hybridization mixture, which consists of:

For rU$_{17}$	OR	For dT$_{17}$
400 nM rU$_{17}$		200 nM dT$_{17}$
400 nM MB		200 nM MB
20 mM MgCl$_2$		3 mM MgCl$_2$

- Suprasil quartz– synthetic fused silica (SiO$_2$) materials, of high purity, manufactured by the flame hydrolysis of SiCl$_4$. It has excellent physical properties as well as outstanding optical characteristics in the deep UV and the visible wavelength range.

.

Results & Discussion

The lengths of the probe and target sequence were carefully designed to enable the best detection of the damage site in oligonucleotides, using the results of previous studies. The main parameter to be held constant was that the binding strength of the stem for itself should be similar to the binding strength of the loop for the target sequence. This is an important factor to ensure that the MB is efficient in discriminating a single damage site in an oligonucleotide. The MB was synthesized with a loop and a stem. The loop consists of 17 poly-A nucleotides which are complementary to the targets. The stem, attached to the loop at either end, is made up of two sequences each of which consists of a stretch of five

- Parameter – quantity constant in case considered, but varying in different cases.
- Target – the nucleotide sequence to which the MB loop binds.
- 6-FAM – is 6-carboxyfluores-cein.
- Fluorophore - is a component/part of a molecule which causes a molecule to be fluorescent. It is a part in a molecule which will absorb energy of a specific wavelength and re-emit energy at a longer, i.e. lower energy, (but equally specific) wavelength.
- 3-DAB – is 4-(4’-dimethylaminophenylazo)-benzoic acid).

nucleotides which are complementary to each other, and constitute the 5' and 3' ends of the MB. The 5' end of the MB, was labelled by attaching a 6-FAM fluorophore while the 3' end was attached to a 3-DAB quencher. The stem, with its sequence of nucleotides, has no role in binding to the target, but is involved in hairpin formation. If there are damage sites in the target sequence, the probe-target hybrid weakens, thereby favoring reforming the stem helix. There are several factors which influence the performance of the MB, such as:

- Stem length
- G-C content in the stem
- Cation concentration
- Hybridization temperature

dT_{17} and rU_{17} were used as targets because their photochemistry is known very well. When UV-induced damage occurs in dT_{17} the primary photo-products formed are CPD, along with small amounts of the [6-4] pyrimidine-pyrimidinone photoproduct as well; whereas when rU_{17} are subjected to UV-irradiation, under anoxic conditions, CPD and photohydrates are formed in almost equal amounts. The quantum yield (φ) for CPD formation in poly-dT is given by $\varphi = 0.044$, whereas in poly-rU, $\varphi = 0.012$ for photohydrate formation and $\varphi = 0.015$ for CPD formation. Thus by comparing the extent

- Quencher – the part of the MB which accepts electrons, or energy from the fluorophore, causing the fluorescence to be extinguished, or "quenched."
- G-C content – the percentage of the nucleotides guanine and cytosine in the DNA.
- Hybridization – the process of forming a hybrid between a nucleic acid and its complement.
- Hybrid – a strand of DNA bound to its complementary strand to form a double-stranded double helix.
- CPD – Cyclobutane pyrimidine dimers (CPD) are made of two thymine molecules, connected together in such a way as to form a cyclobutane ring between them.
- Pyrimidine-pyrimidinone [6-4] photopro-duct – a minor photoproduct.
- Anoxic conditions – Conditions in which there is no oxygen.
- Photohydrates - are the products formed between water and the pyrimidine bases in DNA and RNA.
- The quantum yield (φ) – of a radiation-induced process is the number of times that a defined event occurs per photon absorbed by the system. The quantum yield is a measure of the efficiency with which absorbed light produces some effect, and may be expressed as:

$$\varphi = \frac{\text{\# of destroyed molecules}}{\text{\# of photons absorbed by system}}$$

of hybridization of MB with rU and dT, which have been photodamaged using UV light, the relative destabilizing effect of the CPD and photohydrates on duplex formation can be determined. Using poly-dT and poly-rU oligonucleotides in this fashion, effectively eliminates complications that may arise due to the occurrence of other processes, such as spontaneous depurination, which requires heat.

Factors affecting MB Hybridization

It has been well established that salt-dependent electrostatic effects play an important role in determining the stability, structure and binding behavior of nucleic acid probes. The studies carried out in this area previously, have used:

1 mM $MgCl_2$ - for DNA-MB studies, and

5-60mM $MgCl_2$ - for RNA-MB hybrids.

In order to find a more optimal/suitable range for the RNA-MB hybrids, this study initially investigated the influence of the Mg^{2+} concentration on the fluorescence intensity of the MB alone, and then of the MB in the presence of the rU_{17} target. The results of this initial investigation, as shown in

Fig. 1, indicated that the fluorescence intensity gradually increased as the concentration of the Mg^{2+} was raised from 0 to 5 mM and does not change much when the

- Depurination – is a DNA alteration in which the hydrolysis of a purine base (i.e. Adenine or Guanine) from the deoxyribose-phosphate backbone occurs. After a depurination, the sugar phosphate backbone remains and the sugar ring has a hydroxyl (-OH) group in the place of the Adenine or Guanine.

- Electrostatic effects – effects due to positive and negative electrical charges interacting with each other.

Mg^{2+} concentration is increased beyond 5mM.

Consequently, the Mg^{2+} concentrations chosen for the rest of the study were:

1 mM for DNA-MB hybrids

5-10 mM for RNA-MB hybrids.

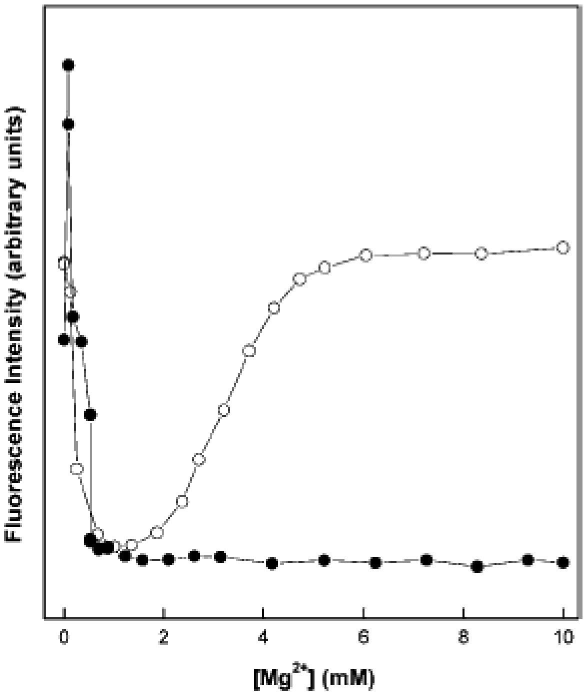


Figure 1. Fluorescence intensity at 520 nm of 210 n*M* MB in the absence (●) and presence (o) of target 2.10 μ*M* rU_{17} as a function of magnesium chloride concentration. Fluorescence excitation wavelength was 492 nm, and the spectra were recorded at room temperature.

In order to make sure that the MB probes can ideally distinguish single damage sites in oligonucleotides, the following conditions have to be met. When the target nucleotides are absent, the MB stem must be closed and must then exhibit minimal fluorescence. In the presence of undamaged target nucleotides, rU_{17} and dT_{17}, the MB must form a stable probe-target hybrid, and exhibit maximum fluorescence. In addition, when a single damage site is present in the target, the binding should be sufficiently destabilized, such that the hairpin configuration has a much lower free energy than the hybridized probe. This condition will make sure that only minimum fluorescence is emitted from the damaged oligonucleotide-MB hybrid, and will thereby ensure a maximum difference in fluorescence between the damaged and undamaged oligonucleotides. Measuring the melting curves for the MB alone, followed by MB in the presence of the oligonucleotide targets can be used to determine the relative binding strengths.

In Fig. 2, which depicts the thermal melting curve of the MB alone, it is evident that the MB is in a closed state, at lower temperatures, and as such the fluorescence is quenched/minimal.

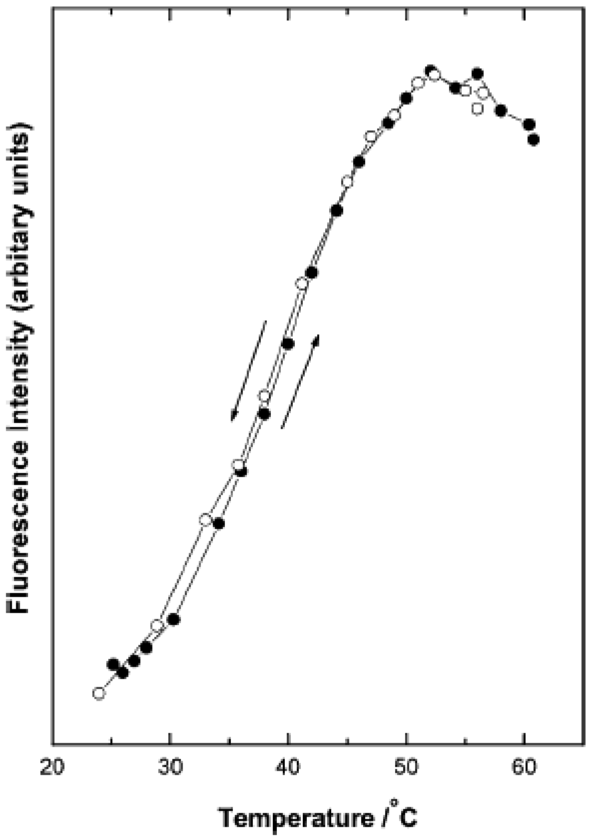


Figure 2. Heating (●) and cooling (o) curves of the 210 n*M* MB alone in buffer containing 1.0 m*M* magnesium chloride. Both curves were generated at a rate of 1.0^{0}C/min.

As the temperature is increased, however, the complementary base pairing is denatured, the stem adopts a random-coil conformation, and consequently, the fluorescence is somewhat restored. The melting point of the MB is approximately 40^{0}C. When the MB was present with a 10-fold excess of the target oligonucleotides (i.e. those which were complementary to the loop sequence in the MB), the melting points are still around 40^{0}C for both rU_{17} and dT_{17}. Actually, the melting point of rU_{17} is slightly lower than that for dT_{17}. In the presence of target oligonucleotides, however, the melting curves are reversed, i.e. at low temperatures the MB fluoresces brightly, but as the temperature is slowly raised, the fluorescence diminishes significantly. When the target oligonucleotides are present, the highly fluorescent MB-oligonucleotide duplex hybrid forms spontaneously at low temperatures. At this point, the fluorescence is maximum, as the quencher and the fluorophore are at their greatest separation. As the temperature is raised, however, the probe-target duplex become destabilized, the MB forms its random coil structure, and the fluorescence is reduced. The transitions described above start at around 28-29^{0}C for rU_{17} and at approximately 30^{0}C for dT_{17}. In this study, therefore, these hybridization temperatures have been selected for the molecular detection of CPD and photohydrate formation in rU_{17} and dT_{17}. The hybrids formed by MB demonstrate differing sensitivities to the different types of phosphoproducts formed. For example, the hybrids are less stable in the presence of photohydrates than CPDs.

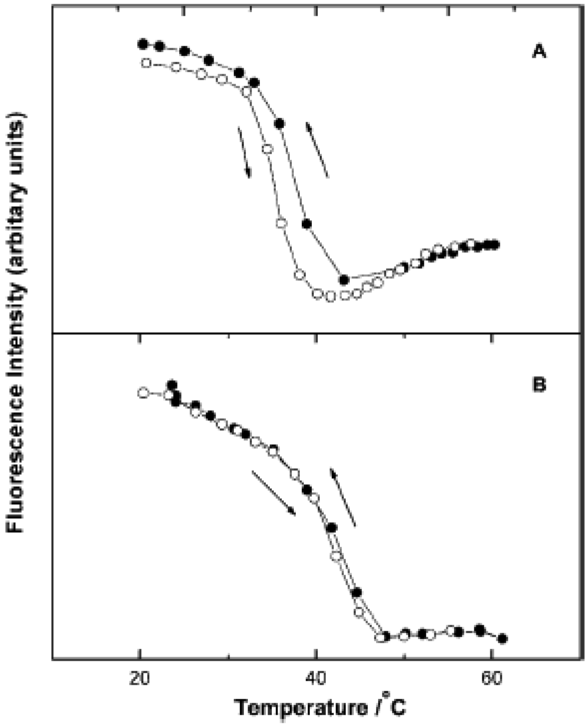


Figure 3. Heating (●) and cooling (o) curves of the 210 n*M* MB in the presence of target (A) 2.00 μ*M* rU_{17} and (B) 2.10 μ*M* dT_{17}. Both curves were generated at a rate of 1.0^{0}C/min.

It is evident from Figs. 2 and 3 that the dissociation of the MB stem occurs at a temperature slightly higher than that of the probe-target duplex: this is an important criterion for the best performance of the MB in the detection of CPD. Consequently, the stem is in a closed state and only unwinds in the presence of the target oligonucleotides, at the hybridization temperature used in this study.

MB detection of CPD damage in oligonucleotides

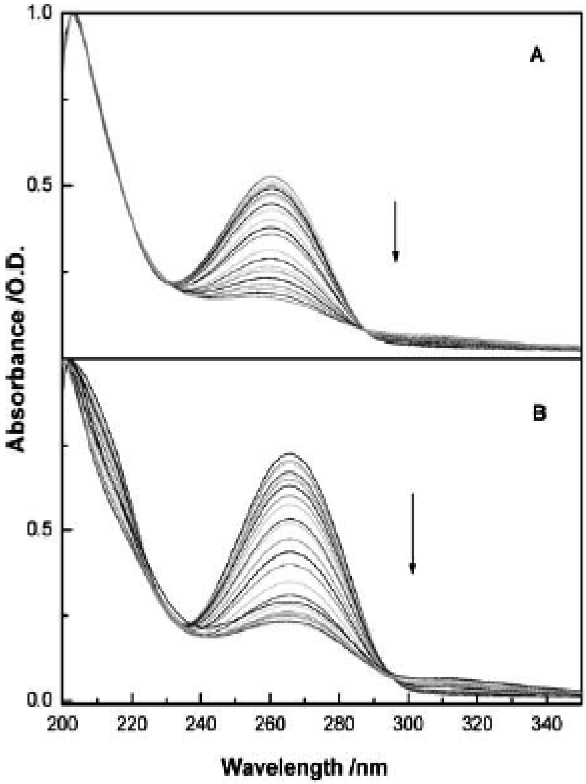


Figure 4. Absorption spectra of 6.67 μ*M* rU_{17} (A) and 6.67 μ*M* dT_{17} (B) in buffer as a function of time of irradiation at 280 nm. Arrows indicate the direction of absorption changes with increasing irradiation time from 0 to 480 min.

Fig. 4 is an illustration of the absorption spectra of rU_{17} and $dT_{17,}$ when irradiated with monochromatic UV light, as a function of time of irradiation at 280 nm. For both the oligonucleotides utilized in this study, the intensity of the 260 nm absorption band gradually decreases/bleaches with increasing irradiation time. This bleaching can be explained in terms of CPD formation in dT_{17}, and with

- Monochromatic light – is light of one color, having wavelengths confined to an extremely narrow range.

regard to CPD and photohydrate formation in rU_{17}, along with the resultant loss of the C_5=C_6 bond in all photoproducts. When controls were run in the absence of the irradiation, this bleaching was not observed: this demonstrates that the absorption change observed arose from the irradiation of the samples and the subsequent photochemistry, rather than from any thermal reactions.

In addition, to the bleaching of the 260 nm absorption band, both rU_{17} and dT_{17} demonstrate an isosbestic point at approximately 290 nm, with a slight increase in absorbance around 310 nm, probably due to the formation of very small amounts of the [6-4] photoproduct.

Fig. 5 is a depiction of the MB fluorescence spectra as a function of irradiation time. Spectra were excited at 480 nm, which is the maximum of the 6-FAM fluorophore absorption spectrum. No shift was observed in the fluorescence of the band maximum, and there was no change in the excitation spectrum of the MB as the irradiation time was increased. It is evident from Fig. 5, that the MB showed maximum fluorescence, when it is bound to unirradiated rU_{17} and dT_{17}. The fluorescence, however, was found to decrease with UV irradiation, and continued to quench/decrease with increasing irradiation. It is noteworthy, that within 50 min of irradiation, the fluorescence reaches that of the background signal of the unhybridized MB.

- Isosbestic point – This term is usually employed with reference to a set of absorption spectra plotted on the same chart for a set of solutions in which the sum of the concentrations of two principal absorbing components, A and B, is constant. The curves of absorbance against wavelength (or frequency) for such a set of mixtures often all intersect at one or more wavelengths, called isosbestic points.

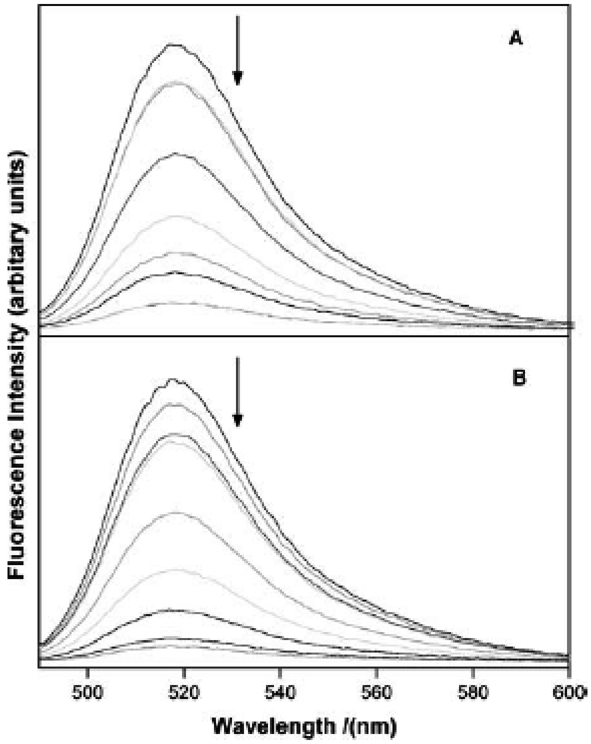


Figure 5. Fluorescence spectra of MB excited at 480 nm as a function of irradiation time at 280 nm. Direction of arrow indicates the direction of the fluorescence intensity change with increasing time of irradiation from 0 to 100 min for (A)400 n*M* MB, 400 n*M* of rU_{17} and 20 m*M* of $MgCl_2$ in buffer and (B) 200 n*M* MB, 200 n*M* of dT_{17} and 3 mM of $MgCl_2$ in buffer. Fluorescence spectra were taken after 0, 1, 2, 10, 18, 25, 40, 60 and 100min for (A) and (B).

This fluorescence quenching clearly indicates that the MB is effective and sensitive in detecting the UV-induced CPD and photohydrate formation in the oligonucleotides used in this study.

- Exponential decay – the decrease of some physical quantity according to the exponential law:

  $N(t) = N_0e^{-\lambda t}$ where,

  $N(t)$ = the quantity at time t,
  N_0 = the initial quantity, the quantity at $t = 0$
  t = time
  λ = the decay constant which is a positive number
  e = base of the exponential

- Single-exponential decay – fitting was done using the equation:

  $y(t) = y_0 + A_1e^{-t_1/\tau}$ where,

  y = the absorbance/fluorescence intensity
  y_0 = a constant
  A_1 = preexponential factor
  τ = mean lifetime
  t = variable

- Double-exponential decay – fitting was done using the equation:

  $Y(t) = y_0 + A_1e^{-t/\tau 1} + A_2e^{-t/\tau 2}$ where,

  y = the absorbance/fluorescence intensity
  y_0 = a constant
  A_1 = preexponential factor 1
  A_2 = preexponential factor 2
  τ_1 = mean lifetime 1
  τ_2 = mean lifetime 2
  t = variable

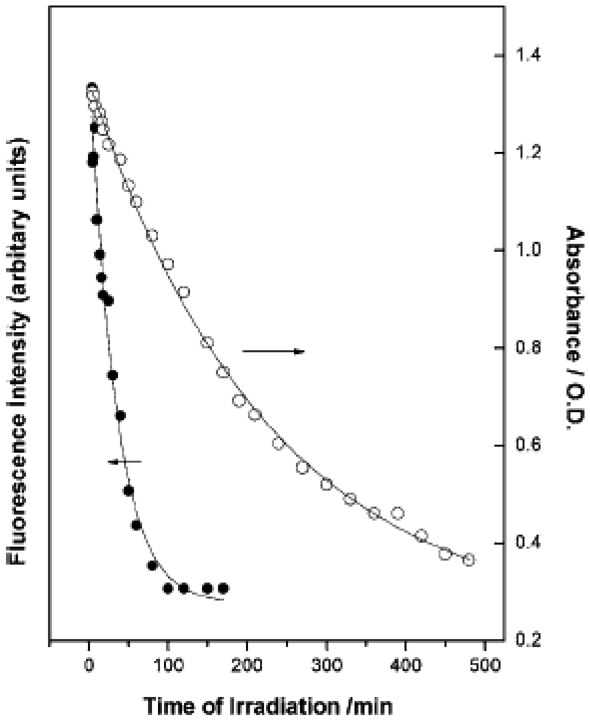


Figure 6. Absorbance (o) of 6.67 µ*M* rU_{17} in buffer monitored at 260 nm and fluorescence intensity (●) at 520 nm of MB monitored as a function of irradiation time. Fluorescence decay data obtained by exciting the hybridization mixture of 400 n*M* rU_{17}, 400 n*M* MB and 20 m*M* $MgCl_2$ at 480 nm. The solid lines through the points are the single-exponential fits resulting in decay times of 123 ± 1 min and 19 ± 0.5 min for absorption and fluorescence, respectively.

The rapid drop in fluorescence to background fluorescence levels suggests that the MB detection of damage has a limited, dynamic range, i.e. the MB can only detect whether the target oligonucleotide is damaged or not, but cannot identify how many lesion sites the target has.

The sensitivity of the MB to damage sites in the target oligonucleotides can be monitored by plotting the fluorescence intensity as a function of irradiation time, and comparing this with a similar plot of the absorption bleach at 260 nm.

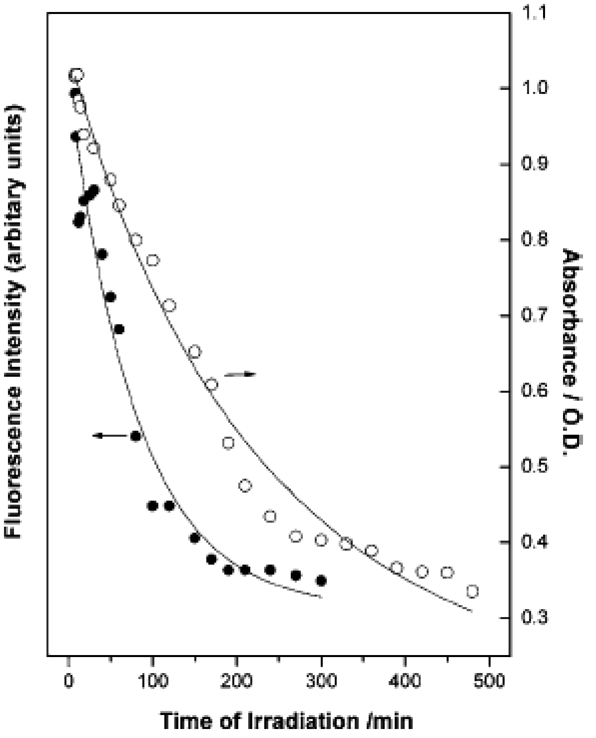


Figure 7. Absorbance (●) of 6.67 μ*M* dT_{17} in buffer monitored at 266 nm and fluorescence intensity (o) at 520 nm of MB monitored as a function of irradiation time. Fluorescence decay data obtained by exciting the hybridization mixture of 200 n*M* dT_{17}, 200 n*M* MB and 3 mM $MgCl_2$ at 480 nm. The solid lines through the points are the double-exponentialfit resulting in decay times of 78 ± 0.5 min (99%) and 180 ± 5 min (1%) for absorption and the single-exponential fit resulting in a decay time of 26 ± 0.5 min for

These results are shown in Fig. 6 for rU_{17}, and Fig. 7 for dT_{17}. As observed from these two figures, the fluorescence always decays faster than the absorption. These results demonstrate how sensitive the fluorescent signal is to the UV-induced damage to the oligonucleotides. Besides, in order to quantify the kinetics, in these Figs. 6 and 7, the absorption and fluorescence data have been fit to either single- or double-exponential decays.

The time constants of these decays along with their amplitude are shown in Table 2. It is evident from Table 2, that the fluorescence intensity for dT_{17}, decays three times faster than its absorption decay: this is significantly lower than the 8.5 times value expected for detection of a single CPD formation site. The formation of a single CPD in dT_{17}, leads to the bleaching of two nucleobases out of the 17. If the MB is capable of detecting a single damage site on each oligonucleotide, it should, on the average, decay 8.5 times faster than the absorption. There are a number of factors that influence the relative rates of absorption and fluorescence decay, perhaps including factors such as the nature of the damage, and the location of the damage. In spite of the fact that the above-mentioned factors are not important in the base-pair mismatch sensitivity of MB, they may be more significant in the detection of CPD and other damage. Therefore, the slower fluorescence decay observed here, suggests that the damaged dT_{17}-MB duplex is more stable than a single base-pair mismatch.

The fluorescence decay for rU_{17} which is 6.5 times faster than the absorption decay, is only half the expected 12.2

times. If photohydrate formation alone occurred exclusively in rU_{17}, this would lead to a fluorescence decay 17 times faster than the absorption. The expected rate of fluorescence decay is obtained by taking a quantum yield-weighted sum of the decay enhancements expected from photohydrate and CPD formation. The assumption made here is that the formation kinetics are identical for each photoproduct. Since the fluorescence decay in rU_{17} is faster than in dT_{17}, this suggests that either the uracil CPD or the photohydrate (or both) destabilizes the duplex more than the thymine photoproducts.

On taking a closer look at Table 2, it is evident that rU_{17} absorption decays 1.6 times slower than that of dT_{17}, whereas the fluorescence signal of the rU_{17}-MB hybrid decays approximately 1.3 times faster than that of the dT_{17}-MB duplex. These ratios were reproducible even when the irradiation conditions were varied, although the time constants themselves changed with the light fluxes in accordance with the expectation. Similar results were obtained with the broadband irradiation as well, as shown in Table 2. The ratio of absorption decay is identical to the expected ratio of total photoproduct quantum yields for dT_{17}, which is (0.044), and for rU_{17}, which is (0.027). This finding implies that the kinetics of photohydrate formation and CPD formation in rU_{17} are similar, and also that the CPD photoreversion rates in the two oligonucleotides are similar, under the conditions used in this study. It must be borne in mind, however, that the photoreversion rate of poly-rU CPD has never been measured.

The fact that the fluorescence and absorption decay rates, for dT_{17} and rU_{17}, are different suggest that there is a difference in the stability of the damaged oligonucleotide-MB duplex, depending on the type of photodamage induced. These differences in the stabilities of the two oligonucleotides could be attributed to the differences in the conformation, or base stacking interactions of rU_{17} and dT_{17}. There are increased base-base stacking interactions induced by the methyl group, and more ordered structures, present more in poly-dT than in poly-rU. This suggests that CPD formation is less disruptive in dT_{17} than in rU_{17}, as observed in this study, by the lower ratio of fluorescence to absorption decays for dT_{17}. It has indeed been shown, that the formation of a thymine CPD, within a dT_n tract, bends the duplex DNA only by 7^0. More work is being carried out to explore the hybridization of the MB with oligonucleotides containing single, well-defined photohydrates, CPD and [6-4] pyrimidine-pyrimidinones, in order to determine the

destabilizing effect that they impart on duplex formation.	

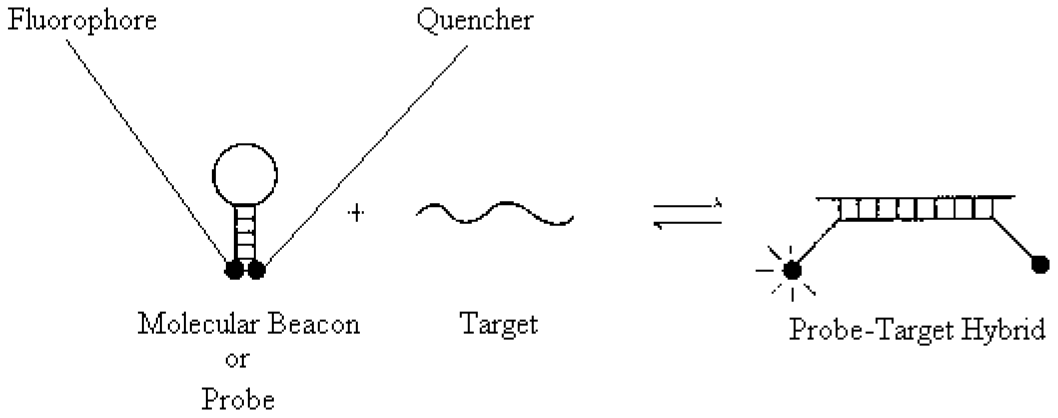
Conclusions

The results displayed in this article indicate how MBs have been used as a probe of CPD and photohydrate damage in oligonucleotides. The methods employed in this study open up a way for the detection and quantitative characterization of damage sites produced as a result of exposure to UV light, ionizing radiation, γ-irradiation, or other means of DNA damage. Using the conditions employed in this study, MB can be efficiently used to detect the UV damage in the oligonucleotides. It appears from this study, that MB are especially sensitive to photohydrate damage in poly-rU, while the greater stability displayed by the duplexes containing CPD, results in a somewhat less sensitive detection of damage.	• MBs – Molecular Beacons • CPD - Cyclobutane pyrimidine dimers

APPENDIX A

Figure 1

Hairpin (Stem-loop) Structure of Molecular Beacons (MB)



Appendix B

Notification of Approval - Delegated Review

Study ID:	Pro00012614	
Study Title:	Using Adapted Primary Literature to teach and test the understanding of Concepts of Evidence in Chemistry held by high school and university undergraduate students	
Study Investigator:	Elizabeth Vergis	
	Approval Date	Approved Document
Date of Informed Consent:	2/22/2010	Initial Information Sheet and Consent Form - Students

Thank you for submitting the above ethics application to the Education, Extension, Augustana and Campus Saint-Jean Research Ethics Board (EEASJ REB). Jerrold Kachur has reviewed your application and, on behalf of the EEASJ REB, approved it as of February 22, 2010. The approval will expire on February 21, 2011.

A renewal report must be submitted prior to the expiry of this approval if your study still requires ethics approval at that time. If you do not renew before the renewal expiry date, you will have to re-submit an ethics application.

Sincerely,

Dr. Stanley Varnhagen, Ph.D.
Chair, Education, Extension, Augustana and Campus Saint-Jean Board (EEASJ REB)

Note: This correspondence includes an electronic signature (validation and approval via an online system).

Appendix B

Participant Consent Form - Students

University of Alberta

I acknowledge that the research procedures have been adequately described, and that any questions I have asked have been answered to my satisfaction. In addition, I know that I may contact the person designated on this form if I have further questions either now or in the future. I have been assured that personal records relating to this study will be kept anonymous. I understand that I am free to withdraw from the study at any time and I will not be asked to provide a reason, but that any data I provide will not be able to be withdrawn.

I agree to the use of my questionnaires for evaluating the understanding of Concepts of Evidence. YES ☐ NO ☐

I agree to be contacted for participation in interviews. YES ☐ NO ☐

If you have agreed to be contacted for participation, please provide your email address on the line below:

____

Please **sign and date** below indicating your willingness to participate in an interview.

______________________ ______________________

(Date) **(Signature of Participant)**

______________________ ______________________

(Date) **(Signature of Investigator)**

Appendix C

Evidence Survey 2010 - Demographics

Please note that all responses will be kept strictly confidential. Thank you for your time and participation!

1. Year of birth ______________________________
2. Were you born in Canada?

   1 Yes Go to 5

   2 No Go to 3

3. If not born in Canada, how long have you lived here? (years) ___________________________
4. If not born in Canada, which country did you come from? ________________________
5. Place of birth (city/town, country) ____
6. How would you describe yourself?

   a. Male

   b. Transgender

   c. Female

   d. Other __________________

7. Canadians come from many cultural or racial backgrounds. Are you **(Please check all that apply)**:

   - ○ Aboriginal, that is North American Indian, Métis or Inuit?
   - ○ Black? (e.g. African, Haitian, Jamaican, Somali)
   - ○ White?
   - ○ Latin American?
   - ○ Chinese?
   - ○ Filipino?
   - ○ Japanese?
   - ○ Korean?
   - ○ Southeast Asian? (e.g. Vietnamese, Cambodian, Indonesian)
   - ○ South Asian? (e.g. East Indian, Pakistani, Punjabi, Sri Lankan)
   - ○ Arab? (e.g. Afghan, Iranian, Turk)
   - ○ Other, specify ____
   - ○ Don’t know
   - ○ Refused

8. I was home-schooled?

   Yes ☐ No ☐

9. I completed Grade 12 in the year _________________

10. Where did you attend High School? (Name, City and Province)

    ____

11. My Grade 12 average was approximately _______%

12. What was the total enrollment of the High School you attended? _______________________

13. The community in which I attended High School has a population of ___________________

14. Have you had any post-secondary educational experience? Yes ☐ No ☐ If yes, please explain below what it was, and in which institution.

    ____

    ____

15. What is the highest level of education of your parent(s) **or** guardian. Pick the person who has had most influence on you. ***(Please circle the number that applies).***

	Mother	Father	Guardian
Less than high school	1	1	1
High school diploma	2	2	2
Trade or vocational (certificate, Diploma, apprenticeship)	3	3	3
Some college	4	4	4
College diploma	5	5	5
Some university	6	6	6
University degree	7	7	7
Graduate degree	8	8	8

16. What is your mother's/guardian's occupation?

____

17. What is your father's/guardian's occupation?

____

18. In what country was your mother born?

____

19. In what country was your father born?

____

20. What language is usually spoken in your family home?

____

21. On average, a family in Edmonton earned approximately $72,000 per year in 2005. Compared to this average, how would you describe your family's financial situation? ***(Please circle the number that applies).***

1 Below average
2 Average
3 Above average
4 Don't know

Thank you very much for taking the time to complete this survey! Please remember that all the information you provided will be kept completely confidential

Appendix D

PRE-TEST

From your reading of Paper 1, focusing on Figure 2, which is below. Please answer the following questions on the lines/spaces provided.

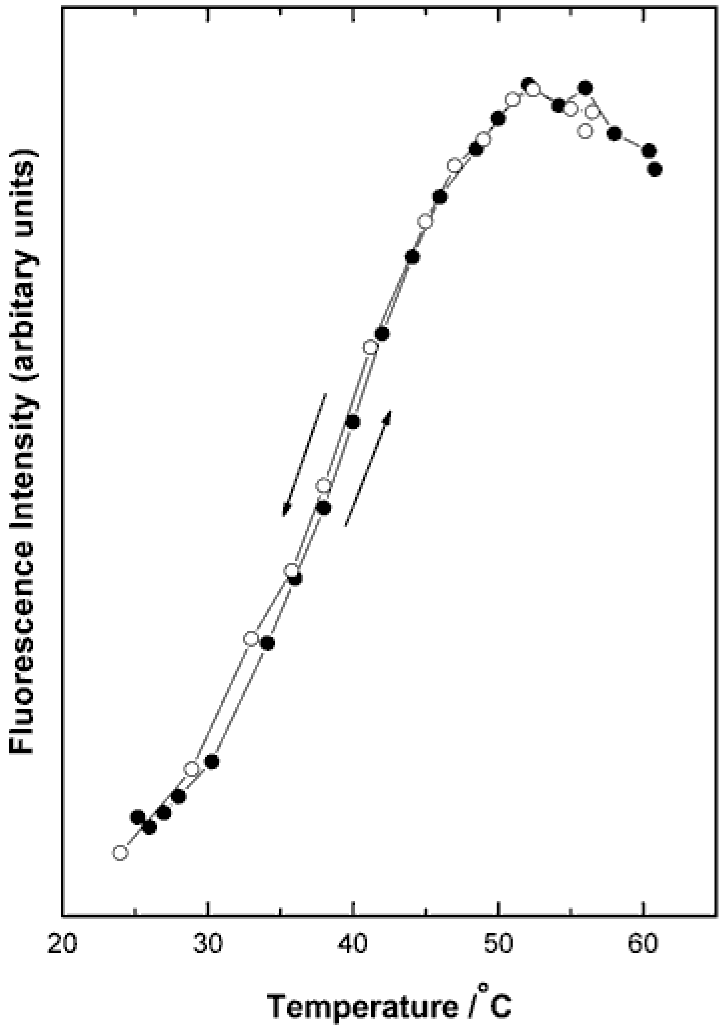


Fig. 2

Figure 2. Heating (●) and cooling (o) curves of the 210 nM MB alone in buffer containing 1.0 mM magnesium chloride. Both curves were generated at a rate of 1.00C/min.

Split Page Here

Q1

1. From the heating curve of the Molecular Beacons (MB) alone in Figure 2, how does the fluorescence intensity change as the temperature is increased?

Q2

2. In what state/conformation are MBs at lower temperatures?

Q3

3. The evidence suggests that the fluorescence of the MBs is low at lower temperatures. Why is this so?

Q4

4. Based on empirical evidence and on molecular structure, what happens to the MBs as the temperature is increased?

Q5

5. In the experiment shown as Figure 2, what is the independent (manipulated) variable and what is the dependent (responding) variable?

Q6

6. What is the name, of the type of calibration that would be needed for the spectrofluorimeter, if the needle at rest, does not point to the zero of the scale?

Q7

7. What is one of the controlled variables in the experiment communicated by Figure 2? Select the answer of your choice.

○ 510 nM Molecular Beacons (MB)

○ 2.10 µM rU17

○ 1.0 mM magnesium chloride

○ 400 nM dT17

Q8

8. Is the experiment depicted in Figure 2, a “fair test”?

○ Yes

○ No

Q9

9. In the question 8, explain your answer. Please be as detailed as possible.

Q10

10. To ensure that a pattern is detected in the fluorescence intensity, what is the range of temperature used in the experiment depicted in Figure 2?
_____^{0}C to ___ ^{0}C.

Q11

11. Is the relationship between fluorescence intensity and temperature shown in Figure 2, an empirical or a mathematical relationship? Indicate your answer by ticking one of the answers given below.

○ Empirical relationship

○ Mathematical relationship

Appendix D

POST-TEST

From your reading of Paper 1, focusing on Figures 6 and 7, which are attached to the back of this questionnaire, please answer the following questions, on the lines/spaces provided below.

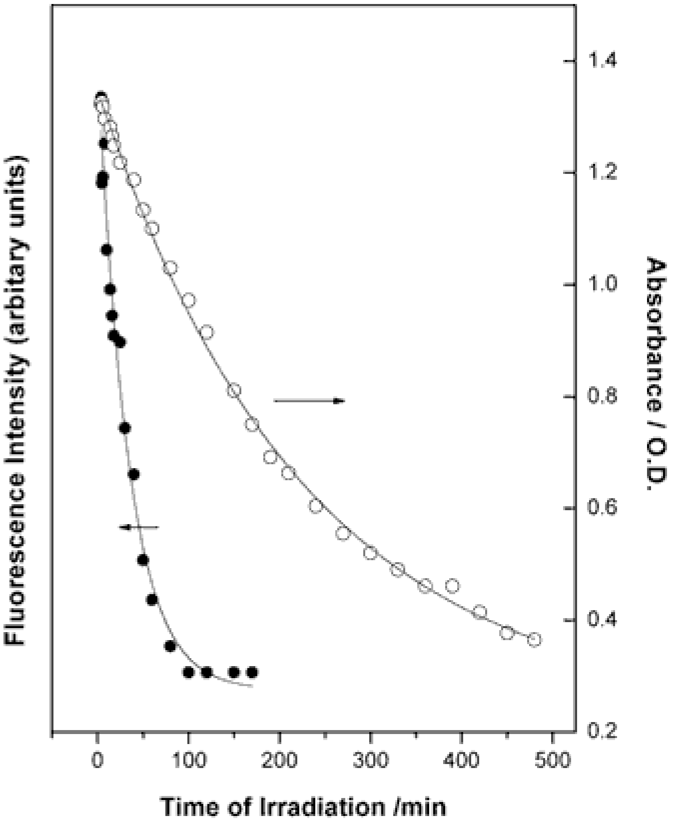


Fig. 6

Figure 6. Absorbance (o) of 6.67 μM rU17 in buffer monitored at 260 nm and fluorescence intensity (●) at 520 nm of MB monitored as a function of irradiation time. Fluorescence decay data obtained by exciting the hybridization mixture of 400 nM rU17, 400 nM MB and 20 mM MgCl2 at 480 nm. The solid lines through the points are the single-exponential fits resulting in decay times of 123 + 1 min and 19 + 0.5 min for absorption and fluorescence respectively.

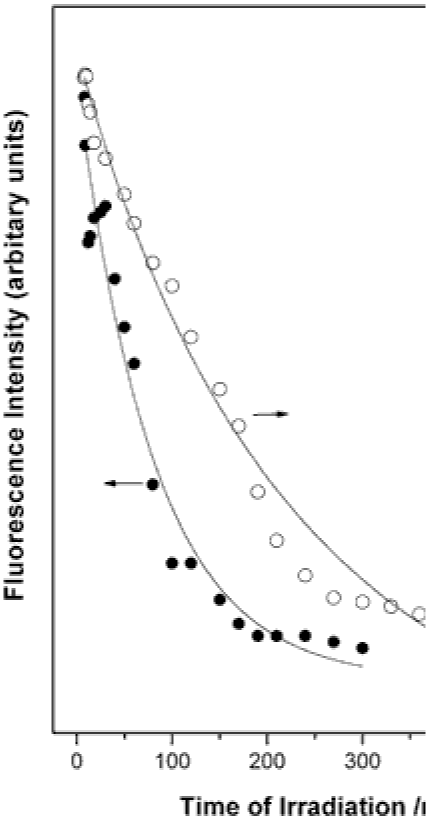


Fig. 7

Figure 7. Absorbance (o) of 6.67 μM dT17 in buffer mo
rescence intensity (●) at 520 nm of MB monitored as a
time. Fluorescence decay data obtained by exciting th
200 nM dT17, 200 nM MB and 3 mM MgCl2 at 480 nm
the points are the double-exponential fit resulting in c
(99%) and 180 + 5 min (1%) for absorption and the si
ing in a decay time of 26 + 0.5 min for fluorescence.

Q1

1. Taking into consideration both Fig. 6 and Fig. 7, in your opinion, which decays faster? (Indicate your choice by one of the answers shown below):

○ The absorption

○ The fluorescence

Q2

2. What does the answer you provided to Question 1 above, tell you about the fluorescence signal?

Q3

3. What is the major difference in the contents of the cuvettes in the experiments depicted in Figures 6 and 7?

	No	**Yes**	**Don't Know**
Cuvette in Fig. 6 has dT17	○	○	○
Cuvette in Fig. 7 has rU17	○	○	○
Cuvette in Fig. 6 has rU17	○	○	○
Cuvette in Fig. 7 has dT17	○		○

Q4

4. What mathematical approximation was done in order to quantify the kinetics in the experiments in Figures 6 and 7?

Q5

5. In the experiment represented as Fig. 6 what is/are the independent variable(s) and what is/are the dependent variable(s)?

	No	**Yes**	**Don't Know**
The dependent variable is absorbance	○	○	○
The dependent variable is fluorescence intensity	○	○	○
The independent variable is	○	○	○

absorbance			
The independent variable is time of irradiation	○	○	○

Q6

6. In the experiment in Fig. 7, describe what happens to the absorbance as the time of irradiation is increased from 0 to 500 min.?

Q7

7. According to the experiment depicted in Figure 6, the limits of detection in absorbance, that can be measured using this spectrophotometer are 0.2 to 1.4 Absorbance Units. If one had a Solution X, with an absorbance higher than the upper limit of detection of this spectrophotometer, could one use this particular spectrophotometer to measure the absorbance of Solution X? Indicate your choice with a tick on the appropriate line.

○ Yes

○ No

Q8

8. (If you answered yes for question 7 above) what is the best method to get around this problem and measure the absorbance of Solution X? Please select the answer you choose below.

○ Prepare a standard curve of absorbance against concentration using different concentrations of Solution X.

○ Dilute Solution X to a range within the standard curve of absorbance against concentration, and note the dilution factor.

○ Using the spectrophotometer, measure the absorbance of the diluted Solution X.

○ Multiply the absorbance of diluted Solution X by the dilution factor.

○ Carry out the steps A to E, in that order, and obtain the absorbance of Solution X.

○ Adjust for the zero error of the spectrophotometer, and measure the absorbance of Solution X.

Q9

9. How have other variables been controlled in the experiment in Fig. 7?

Q10

10. Is the experiment depicted in Fig. 7, a “fair test”?

○ No

○ Yes

Q11

11. For previous question, if you answered yes or no - explain Why? Give reasons for the answer you chose.

Q12

12. Is the relationship between fluorescence intensity/absorbance and time of irradiation shown in Figure 6, an empirical or a mathematical relationship? Indicate your choice by selecting one answer below.

○ Empirical relationship

○ Mathematical relationship

Q13

13. For the question above, give reasons for the answer you chose. Please be as detailed as possible.

Appendix E

Interview Questions for Students

1. In Figure 2 of the APL (Adapted Primary Literature), what do the arrows depict/represent?

   The upward arrow depicts the heating curve as the temperature is raised from 20^0*C to 60*^{0}*C.*

2. What would you do with an anomalous datum in Figure 2 of the APL? Why? Defend/Justify your answer.

   2 things to consider:

   - *Is there a cause/explanation (such as poor measurement procedure) for the aberrant result? If yes, DISCARD result.*
   - *If there is NO cause/explanation more similar measurements have to be taken to see if the anomaly was part of an inherent variation.*

3. Is there sampling in the experimental set up of Figure 6 in the APL? What is sampling? Why do you say that there is/there isn't sampling?

   Yes. Each experiment is a sample out of the total probable batch

.

4. How do we know that the fluorescence intensity is only due to the MBs in Figure 7? What could you do to show that the fluorescence intensity is only due to the MBs and not anything else? How can you set up an experiment to show that the FI is only due to the MBs? What do you call such side-experiments which are set up to shed light on the main experiment?

   Because of the Control done with unirradiated dinucleotides. Control solutions were handled identically but not exposed to the UV light source.

5. What is the scientific purpose of the experiments in the APL as a whole? Why did they do this whole series of experiments from Figs 1 – 7?

   To find the optimal conditions for and to compare the efficacy of Molecular Beacon Probes in measuring the Photodamage in Thymine and Uracil Oligonucleotides, when exposed to UV light

6. Did you enjoy reading the APL? Why? Which part of it was most exciting or interesting for you?

7. Did you notice about the genre used in the APL? What did you notice? Was this done intentionally? What do you think?

   Paper 1 = APL - Argumentative genre

8. In the evidence Survey 2011, what were the **positive aspects** and what were the **negative**

aspects?

9. Of all the CoEs you came across, which are the most important, do you think? Why?

   Design: Variable Structure - CoE 11
   Instruments: Calibration and Error – CoE 5
   Detecting Patterns – CoE 19
   Design: Validity, "Fair Tests" and Controls – CoE 12

10. What do you mean by a "Fair Test"? Is Fig 6/7 a FT? Why or why not?

    The Dependent Variable is observed and only ONE Independent Variable is varied at a time.

11. What did you think of the Teaching Intervention?

    Was it useful? How?

    Were you able to perform better in the Post-Test? Yes/No?

    Was this due to the Teaching Intervention?

    Did the TI help your understanding or boost your confidence level in going into doing the Post-Test?

Appendix E

Interview Questions for Instructors/Teachers

1. What does the similarity in shape between the heating and cooling curves in Figure 2 of Paper1 tell you about the underlying structure of the MBs? Why are both the heating and cooling curves gathered?

At low temperatures the MBs form this hairpin structure. So the stem is bound together and it's got a free loop, and the stem has at each end a quencher and a fluorophore. So it tells us that the quencher and fluorophore are close together. And then as the temperature increases, we're melting, or breaking hydrogen bonds between bases on the stem and so that allows the fluorophore and quencher to separate, so the fluorescence increases, it is not quenched as much. And so finally it reaches the highest temperatures at sort of constant level, and that tells us that the DNA has adopted a structure we call a random coil, in which the fluorophore and quencher are a characteristic distance apart. And the fact that they are the same, the heating and the cooling curves, tells us that, that process is reversible. So it tells us that the structure goes through this random coil, and then reassembles back into the hairpin structure, without any change in the fluorescence data.

If they weren't the same, that process we call hysteresis, and it would tell us that either the process is not reversible, or that the kinetics of the reversible process are slower than the rate at which we did the heating and the cooling.

2. Why are arbitrary units assigned to the fluorescence intensity in Figs. 1, 2, 3, 5, 6, and 7?

Fluorescence is difficult to quantitate. There are a lot of experimental conditions which are not important. The authors are only interested in the change in fluorescence, ΔF.

Fluorescence in general is difficult to quantitate. And so the units we get for FI would vary differently for somebody else trying to do the same experiment, because of a lot of experimental considerations which aren't important because all we are interested in is a change in FI, but which makes it more accurate, to represent the FI in arbitrary units, because there are so many experimental variables, which the FI depends on but which aren't so important in actually understanding the results.

3. Why was the range of evidence restricted to 25-60^{0}C?

The authors have software to predict where the Melting Temperatures of the MBs will lie/be. They used water as a solvent because it remains a liquid in the range of 5^{0}C to 85^{0}C. The Melting Temperatures also lie within this range.

4. What kind of information about the evidence is not provided by theses graphs in Paper 1, that is sometimes provided on graphs – e.g. error bars? Why?

Error Bars – quite small in these experiments. Order of sizeable units. They were quantitated systematically.

When we designed these MBs, with the Wender Website, they have a software tool that allows us to predict what the melting temperature will be so the temperature at which you are half way in *between the hair pin structure and the random coil structure. And so we choose a temperature range that has that in the centre, and we take enough on either side within the range dictated by our solvent. So we use water, we can't go much below about 5^0C, and we can't go much above 85-90^0C. So we need to keep the Melting Temperatures within those ranges because of that constraint, and we try to choose as large a range within there, that is necessary to see both the hair pin structure and the random coil structure, and their FI levels.*

5. What is the scientific/social/technological purpose of this series of experiments?
<u>Hypothesis</u> *– Authors knew that MBs were sensitive to specific base differences in DNA. The question they wanted answered was:* ***Can MBs select specific damage?***

6. How are the best fit lines created in Figs. 2, 6, 7?

7. What was the criterion used in selecting the target oligonucleotides? Why? Explain your answer?

T & U were most active. Authors wanted Oligonucleotides where they could be sure that damage would occur. They wanted the experiments to be ***<u>predictive</u>***.

8. What is the transition that occurs in rU_{17} at approximately 28-29^0C in fig. 3? Explain clearly the change that occurs at the molecular level at this transition temperature.

9. What is the relationship between empirical and theoretical knowledge in this report?

10. Is the level of reliability and validity of this whole investigation sufficient? How could you increase it?

11. Are CoEs empirical or are they theoretical? What is your opinion?

12. What do you think is the role of CoEs in the teaching of Chemistry? Can the teaching of Chemistry be made more effective by introducing CoEs?

13. How would you increase the participation, or the response rate?

14. What did you think of the Teaching Intervention? Can it affect/change a student's understanding of CoEs? Was it detailed enough about CoEs?

Appendix F

Codebook For Student Interviews

Question 1

a) Arrows depict:

1. Upward arrow - Increase in temperature. Downward arrow – Decrease in temperature
2. Upward arrow – Increase in FI. Downward arrow – Decrease in FI.
3. Upward arrow – Increase in temp & FI. Downward arrow – Decrease in temp & FI.
4. Other
9. Not sure/Don't know
10. Unable to determine

Question 2

a) What do you call data that do not quite obey the general trend?

1. Outliers
2. Anomalous Results
3. Outliers & Anomalous Results
4. Other
9. Not sure
10. Unable to determine

b) Two things you could do with Anomalous Results:

1. Explain the results
2. Repeat the experiment
3. 1 & 2
4. Other
9. Not sure
10. Unable to determine

Question 3

Is there sampling in Fig 6? Why?

1. Yes
2. No
3. 1 or 2 and explanation of why. [For 1 – Each experiment is an aliquot out of the total probable batch/total hybridization mixture] OR [Not taken entire Hybridization Mixture, but only an aliquot]
4. Other
9. Not sure
10. Unable to determine

Question 4 – This question with all its parts refers to Fig 7 in the APL. The hybridization mixture, in the main experiment represented as Fig 7, is 200 nM dT_{17}, 200 nM MB and 3 mM

$MgCl_2$. The Absorbance (o), and Fluorescence Intensity (FI) (•), of the hybridization mixture, are monitored as a function of Irradiation Time. Please answer the following questions based on Fig 7.

a) What do you call experiments that we do on the side, to shed light on/strengthen the main experiment?

1. Control Experiments
2. Not sure – finally got Control Experiments
3. Other
9. Not sure
10. Unable to determine

b) Referring to 6.6 µM dT_{17}, what is T?

1. Thymine/Uracil
2. Other
9. Not sure
10. Unable to determine

c) What does the 'd' stand for?

1. DNA/RNA
2. Other
9. Not sure/No answer
10. Unable to determine

d) How could you set up a Control Experiment to show that the FI observed in this experiment was really due to the MBs and not anything else?

1. Having a Control with just MBs
2. Having a Control with just DNA
3. Having a Control with DNA & $MgCl_2$
4. Have Control with DNA & $MgCl_2$ and add MBs to that
5. Other
9. Unsure mixtures of above which are not quite correct
10. Unable to determine

e) The fluorescence of the Control is:

1. High
2. Low
3. Lower than in the main experiment
4. Other
9. Not sure
10. Unable to determine

Question 5

***Scientific purpose of this whole series of experiments**:

1. To find the optimal conditions for and to compare the efficacy of MB probes in measuring the Photodamage in Thymine and Uracil Oligonucleotides when exposed to the UV-light source.
2. Other
9. Not sure
10. Unable to determine

So how much damage are they hoping to detect? --- For Bank entry

1. A single base-pair damage/A single base mispair

<u>Question 6</u>

a) Enjoyed reading the APL:

1. Yes
2. Yes, but .....
3. No
4. No, but .....
5. Other
9. Not sure
10. Unable to determine

b) The most exciting/interesting/useful part of the APL:

1. The annotations and meanings provided in right column
2. Other
3. None
9. Not sure
10. Unable to determine

<u>Question 7</u>

The genre used in the APL was:

1. Descriptive/like an Explanation/Experimental
2. Narrative
3. Argumentative
4. 1 & 2
5. 1 & 3
6. Other
9. Not sure
10. Unable to determine

<u>Question 8</u>

a) The positive aspects in Evidence Survey 2011:

1. New
2. Fun/Exciting/Interesting – not too intense
3. Electronic access – didn't have to be anywhere physically
4. Other

5. Not applicable
9. Not sure
10. Unable to determine

b) The negative aspects in Evidence Survey 2011:

1. TI/Video could hear movements of camera
2. Read and do not understand - discouraging
3. Other
4. Not applicable
9. Not sure
10. Unable to determine

<u>Question 9</u>

a) The most important CoE was:

1. CoE #11 – Design: Variable Structure
2. CoE #5 – Instruments: Calibration and Error
3. CoE #19 – Detecting Patterns
4. CoE #12 – Design: Validity, Fair Tests and Controls
5. Other
9. Not sure
10. Unable to determine

***b) Why?**

<u>Question 10</u>

a) Initial response to defining a Fair Test:

1. Not at all accurate/Minimum accurac
2. Medium accuracy
3. Full accuracy - The DV (may be more than one) is observed and only ONE IV is varied AT A TIME

When probing the levels of understanding involved in a Fair Test:

b) What are being measured in Fig 6?

1. FI
2. Abs
3. FI & Abs
4. Hesitant, but gets FI & Abs finally
5. Other
9. Not sure
10. Unable to determine

c) What sort of variables are FI and Abs?

1. DV

2. Responding
3. IV
4. Manipulated
5. Other
9. Not sure
10. Unable to determine

d) What sort of variable is Time of Irradiation?

1. IV
2. Manipulated
3. DV
4. Responding
5. Other
9. Not sure
10. Unable to determine

e) How many IVs are being manipulated at ONE TIME?

1. Two
2. One
3. Other
9. Not sure
10. Unable to determine

f) When is an experiment in Science valid?

1. Not sure
2. Fairly muddled
3. When only one IV is varied at a time, during the course of the experiment <u>OR</u> when the experiment is a Fair Test.
4. Other
9. Not sure
10. Unable to determine

g) Could you change the Time of Irradiation and the Magnesium Concentration in the same experiment?

1. Yes
2. No
3. Hesitant, but finally states "No".
4. Other
9. Not sure
10. Unable to determine

<u>Question 11</u>

a) What did you think of the Teaching Intervention (TI)?

1. Very useful/Excellent/Pretty useful/Really useful
2. Useful

3. Somewhat useful – with reservations
4. Not Useful
5. Other
9. Not sure
10. Unable to determine

***b) How was it useful/not useful?**

1. Not just repetition of what was in the papers. Had some other related concepts covered as well
2. Drawings – like flow chart
3. Other
9. Not sure
10. Unable to determine

c) Were you able to perform better on the Post-Test?

1. No
2. Yes
3. Other
9. Not sure
10. Unable to determine

d) Was this due to the TI?

1. No
2. Yes
3. Other
9. Not sure
10. Unable to determine

e) Did the TI help your understanding or boost your confidence level in going into doing the Post-Test?

1. Boosted my confidence level
2. Helped me understand better
3. Confidence higher. Understanding better.
4. Other
9. Not sure
10. Unable to determine

		Student ID:	Coder ID:							
		Coding Form	1	2	3	4	5	6	9	10
Q1		Arrows depict								
Q2	a)	Data that do not obey General Trend are called...								
	b)	Two things done with Anomalous Results								
Q3		Sampling in Fig 6? Why?								
Q4	a)	Name side experiments done to shed light on main expt.								
	b)	What is T?/U								
	c)	What is 'd'?/'r'?								
	d)	Side expts done to shed light on main one, composed of ..								
	e)	Fluorescence of Side Experiment above ...								
Q5		Scientific purpose of this whole series of experiments								
Q6	a)	Enjoyed reading the APL?								
	b)	Most interesting part of APL?								
Q7		Genre used in the APL								
Q8	a)	Positive Aspects of Evidence Survey								
	b)	Negative Aspects of Evidence Survey								
Q9	a)	The most important CoE?								
	b)	Why?								
Q10	a)	Intitial response to defining a Fair Test								
	b)	What are being measured/observed in Fig 6?								
	c)	What sort of variables are FI and Abs.?								
	d)	What sort of variable is Time of Irradiation?								
	e)	How many IVs are being manipulated at ONE TIME?								
	f)	When is an experiment in Science valid?								
	g)	Change Time of Irradiation and [MgCl2] in same expt?								
Q11	a)	What did you think of the TI?								
	b)	Howwas it useful/not useful?								
	c)	Able to perform better in Post-Test?								
	d)	Was this due to the TI?								
	e)	Did the TI help your understanding or your confidence?								

Appendix F

Codebook - for Instructors

Question 1

a) Similarity in shape

1. Changes at molecular level – hair-pin at low temps and random coil at high temps
2. Changes are reversible
3. 1 and 2
4. Other
9. Not sure
10. Unable to determine

b) Bothe heating and cooling curves gathered because:

1. It tells us that the process is reversible and there is no hysteresis
2. It tells us that the structure goes through this random coil structure and reassembles
   back into the hairpin structure, without any change in the fluorescence data
3. It tells us that the process is reversible
4. Other
9. Not sure
10. Unable to determine

Question 2

Arbitrary Units assigned to FI because:

1. Fluorescence is difficult to quantitate
2. Authors only interested in a change in FI, ΔF.
3. It doesn't matter. It's on a relative scale. It is relative intensity
4. Other
9. Not sure
10. Unable to determine

Question 3

Why was the range of evidence restricted to 25-60^{0}C

1. The Melting Temperature of the MBs was predicted to be half-way between hair-pin and random coil structures (i.e. ~45^{0}C). Chose a temperature range that has 45^{0}C in the centre, and 20^0 above and 20^0 below (i.e. 25-60^{0}C).
2. Within physiological range of temperature
3. Water chosen as the solvent, therefore the temperature range has to be within 5-90^{0}C
4. Other
9. Not sure
10. Unable to determine

Question 4

What kind of information about evidence is not provided by these graphs, that is sometimes provided – e.g. Error Bars? Why?

1. Error Bars - are quite small, the size of points for the random errors.
2. The fact that the Error Bars are quite small is mentioned in the legends
3. The systematic errors have not been quantitated
4. Other
9. Not sure
10. Unable to determine

Question 5

What is the scientific/social/technological purpose of this series of experiments?

1. Authors wanted to test whether the MBs would detect damage at a base, in the same way that they detect a different base at that site.
2. Authors knew the MBs were sensitive to specific base differences in DNA. Their objective was to find out whether the MBs could detect specific base damage.
3. Their hypothesis was the MBs can be used to indicate that uv-light will damage DNA/RNA.
4. Other
9. Not sure
10. Unable to determine

Question 7

What was the criterion used in selecting the target oligonucleotides? Why? Explain your answer.

1. Authors knew that Thymine & Uracil were the most photochemically-active bases in DNA and RNA respectively.
2. They wanted a very well-defined, simple sequence where they could be sure damage would occur, and where they knew what that damage would be, i.e. they wanted the experiment to be predictive.
3. If a random sequence of nucleobases was used, the possibility of photoproducts increases exponentially.
4. This was the most predictive experiment
5. Other
9. Not sure
10. Unable to determine

Question 9

What is the relationship between empirical and theoretical knowledge in this report?

1. Chemistry uses indirect information to infer what goes on at the molecular level, and has models for the molecular level which are fairly consistent.

2. The empirical evidence is fluorescence and spectroscopic evidence, and a lot of molecular and microscopic inferences can be made based on this evidence, because they are in line/sync with properties inferred from previous experiments.
3. Empirical evidence in this report sheds light on what can be supported as theoretical knowledge, e.g. hair pin and random coil structure of MBs.
4. Theoretically it is strongly suspected that uv-light could damage DNA/RNA. Here you have a case, where there is empirical proof (in the laboratory), that this is actually true.
5. Other
9. Not sure
10. Unable to determine

Question 10

a) Is the level of reliability and validity of this whole investigation sufficient?

1. Yes. This is preliminary work, but has always been consistent, especially structure of MBs and how they work. MBs are pretty reliable in detecting DNA damage.
2. It is a very succinctly put together paper with good figures that show exactly what they are trying to say. Probably just giving, ... may be a little more information on the ..... like in the figures themselves. They do give data for their fitted curves. They give + or – data in terms of how well the curves fit, but .... I thought it was a well done paper.
3. My best guess is yes, it is sufficient. This work, having been published in a peer-reviewed journal, I would assume other people in this field would deem it sufficiently reliable and valid. This is what I would guess, but I cannot myself judge it.
4. Other
9. Not sure
10. Unable to determine

b) How could you increase it?

1. An issue not resolved is: How much damage do you have to have in sequence before the MB can detect that?
2. Sensitive enough to detect one base pair damage
3. Repeating the experiment is one way of increasing the reliability/validity
4. Other
9. Not sure
10. Unable to determine

Question 11

Are CoEs empirical or are they theoretical? What is your opinion?

1. CoEs are empirical
2. CoEs are theoretical
3. Some CoEs are empirical while others are only theoretical
4. Some CoEs are empirical with theoretical implications/underpinnings, while others are theoretical with empirical applications
5. Other
9. Not sure
10. Unable to determine

Question 12

a) What do you think is the role of CoEs in the teaching of Chemistry?

1. To shed more light on the role of Procedural Knowledge in the teaching of Chemistry
2. To stipulate a consideration of the factors controlling the collection, measurement and handling of data
3. To lay down the theory behind the factors to consider when designing/planning an experiment/investigation
4. Other
9. Not sure
10. Unable to determine

b) Can the teaching of Chemistry be made more effective by introducing CoEs?

1. Yes
2. No
3. Does not make any difference
4. Other
9. Not sure
10. Unable to determine

Question 13

How would you increase the participation, or the response rate?

1. Change the time at which the Evidence Survey was done to earlier in the year
2. Change the time at which the Evidence Survey was done to later in the year
3. Give a greater monetary reward
4. Other
9. Not sure
10. Unable to determine

Question 14

a) What did you think of the Teaching Intervention?

1. The concepts are better understood because of the TI
2. The visuals provided by the TI helped students to grasp the concepts better
3. The TI brings it all together
4. Other
9. Not sure
10. Unable to determine

b) Can it affect/change a student's understanding of CoEs?/Do you think a TI like that is effective?

1. Yes, the visuals and anecdotal explanations sure help to change students' understanding of CoEs
2. The first-hand report by one of the authors, the unpacking of the instrumentation used, and the anecdotal references to experiments in the paper, the visual representations of the MBs and their reactions at the molecular level, and the explanations and representations of CoEs using diagrams, do enhance the understanding of CoEs, especially for visual learners
3. It has to, I think, because that is what teaching is
4. Other
9. Not sure
10. Unable to determine

c) Was it detailed enough about CoEs?

1. Yes
2. Yes, but
3. No
4. No, but
5. Other
9. Not sure
10. Unable to determine

		Instructor ID:	Coder ID:							
		Coding Form - Instructors	1	2	3	4	5	6	9	10
Q										
1	a)	Similarity in shape								
	b)	Both heating and cooling curves gathered because								
2		Arbitrary Units assigned to FI because:								
3		Why was the range of evidence restricted to 25-60°C								
4		What kind of info about evidence is not provided by graphs? Why?								
5		Scientific/social/technological purpose of series of expts.								
7	a)	What was Criterion used in selecting target nucleotides?								
	b)	Why? Explain your answer.								
9		Relationship between empirical and theoretical knowledge?								
10	a)	Is level of reliability/validity of this whole investigation sufficient?								
	b)	How could you increase it?								
11		Are CoEs empirical or theoretical?								
12	a)	What do you think is role of CoEs in the teaching of Chemistry?								
	b)	Teaching of Chemistry made more effective by introducing CoEs?								
13		How would you increase the participation/response rate?								
14	a)	What did you think of the TI?								
	b)	Can it affect/change a student's understanding of CoEs?								
	c)	Was it detailed enough about CoEs?								

Appendix G

Categories and Sub-categories of Concepts of Evidence from Gott et al.

Concepts of Evidence	*Degree of Complexity*
1. Fundamental ideas	3
1.1 Opinion and data	
1.2 Links	
1.3 Association and causation	
1.4 Types of measurement	
1.5 External tasks	
2. Observation	1
1.1 Observing objects	
1.2 Observing events	
1.3 Using a key	
1.4 Taxonomies	
1.5 Observation and experiment	
1.6 Observation and map drawing	
3. Measurement	2
4. Instruments: Underlying relationships	2
a. Linear relationships	
b. Non-linear relationships	
c. Complex relationships	
d. Multiple relationships	
5. Instruments: Calibration and error	2
1.1 End points	
1.2 Intervening points	
1.3 Zero errors	
1.4 Overload, limiting sensitivity/limit of detection	
1.5 Sensitivity	
1.6 Resolution and error	
1.7 Specificity	
1.8 Instrument use	
1.9 Human error	
6. Reliability and validity of a single measurement	2
1.1 Reliability of measurements	
1.2 Reliability of instruments	
1.3 Reliability based on human error	
1.4 Validity	
7. Choice of an instrument for measuring a datum	2
8. Sampling a datum	3
9. Statistical treatment of measurements of a single datum	3
10. Reliability and validity of a datum	2

11. Design of investigations: Variable structure 3
    1.1 The independent variable
    1.2 The dependent variable
    1.3 Correlated variables
    1.4 Categoric variables
    1.5 Ordered variables
    1.6 Continuous variables
    1.7 Discrete variables
12. Design: Validity, "fair tests" and controls 3
    1.1 Fair test
    1.2 Control variables in the laboratory
    1.3 Control variables in field studies
    1.4 Control variables in surveys
    1.5 Control group experiments
13. Design: Choosing values 3
14. Design: Accuracy and precision 3
15. Design: Tables 3
16. Reliability and validity of the design 3
17. Data presentation 3
    1.1 Tables
    1.2 Bar charts
    1.3 Line graphs
    1.4 Scatter graphs
    1.5 Histograms
    1.6 Other forms of display
18. Statistics for analysis of data 3
19. Patterns and relationships in data 3
    1.1 Types of patterns
    1.2 Linear relationships
    1.3 Proportional relationships
    1.4 'Predictable' curves
    1.5 Complex curves
    1.6 Empirical relationships
    1.7 Anomalous data
    1.8 Line of best fit
20. Reliability and validity of data in the whole investigation 3
    1.1 A series of experiments
    1.2 Secondary data
    1.3 Triangulation
21. Relevant societal aspects 3

Appendix H

∆diff From Post-Test Score minus Pre-Test Score

Student Number	Pre-Test Total /19	Post-Test Total /19	Post-Test minus Pre-Test ∆diff
Student 1	9.75	10	0.25
Student 2	13.5	13.75	0.25
Student 3	11.75	10	-1.75
Student 4	8	11.25	3.25
Student 5	7.25	8.5	1.25
Student 6	5.5	9	3.5
Student 7	11.75	14.5	2.75
Student 8	5.75	9.5	3.75
Student 9	7.75	9.25	1.5
Student 10	3.75	3	-0.75
Student 11	3.5	6	2.5
Student 12	12.75	9.5	-3.25
Student 13	5	7.25	2.25
Student 14	8.75	11.5	2.75
Student 15	7	10.5	3.5
Student 16	7	7.25	0.25
Student 17	8.5	13.75	5.25
Student 18	6.75	8.5	1.75
Student 19	6.5	11.5	5
Student 20	8	11	3
Student 21	7.75	9.25	1.5
Student 22	7.5	7.75	0.25
Student 23	8.25	14.5	6.25
Student 24	7	8	1
Student 25	11.75	15	3.25
Student 26	9	10	1
Student 27	7.25	8.5	1.25
Student 28	7.75	10.5	2.75
Student 29	10.5	14.75	4.25
Student 30	8.25	9.75	1.5
Student 31	6	11.25	5.25
Student 32	5	9	4

Student 33	9	6	-3
Student 34	10	11.25	1.25
Student 35	7.5	11.25	3.75
Student 36	5.5	11.75	6.25
Student 37	6.25	7.5	1.25
Student 38	8.75	12	3.25
Student 39	13.5	14	0.5
Student 40	8.75	12.25	3.5
Student 41	12	12.75	0.75
Student 42	10.25	11.75	1.5
Student 43	10.75	10.75	0
Student 44	9.5	10.25	0.75
Student 45	4.25	9.25	5
Student 46	9.75	12.5	2.75
Student 47	9.75	11	1.25
Student 48	9.5	6.5	-3
Student 49	11.5	11.75	0.25
Student 50	11	14.5	3.5
Student 51	10.5	9	-1.5
Student 52	11.5	11	-0.5
Student 53	10	14.5	4.5
Student 54	10.25	12	1.75
Student 55	6.75	12	5.25
Student 56	7	11.75	4.75
Student 57	10.25	13.5	3.25
Student 58	10.75	10.25	-0.5
Student 59	9.5	8.5	-1
Student 60	5.75	6.75	1
Student 61	2.5	4	1.5

Appendix I

List of Findings which Constituted the Themes and Ideas from Participants which Contributed to the Findings

Part A: From Student Interviews

Themes	Findings	Ideas from Participants
1Grappling with the APL	Student experiences from reading APL	enjoyed reading the APL, some needed extra help, APL a little dry, some parts of APL hard to understand and read, APL a bit too long, APL was not entertaining but enjoyable academically, valued very highly participating in my study, APL enjoyable at the beginning but boring by the end, liked the annotations and meanings in the APL, the meanings helped a lot, first reading was difficult but got easier as the readings were repeated, very interesting for sure, did not enjoy the APL, preferred Primary Literature to APL, preferred listening to video more than reading the APL, video was easier to understand, APL straightforward and easy to understand, time-consuming, a task where you were on your own, APL took some close critical reading, crash reading once or twice, had the added bonus of the definitions on the side, a little hard to understand even with the annotations, preferred APL to Primary Literature, without the definitions would not know what any of the words meant, could for the most part understand what was going on in APL, reading the APL was an enjoyable experience

Part A: From Student Interviews (Continued)

Themes	Findings	Ideas from Participants
	Intriguing parts of APL	the annotations and meanings provided, the beginning or introduction of APL, applications of this work, APL provided a good background to the study, materials and methods section, explanations on the graphs, molecular beacons could be used to detect what was happening in the DNA, use of spectrophotometry, interdisciplinary nature of the work in the APL, APL was okay, not really what I am interested in, scientific methods used to detect damage in the base-pairs, the light and dark phases of the molecular beacons, enjoyed the graphs comparing the DNA and RNA, found such comparisons most exciting and enjoyable, understanding APL got easier after repeated reading, could figure out damage in base-pairs of DNA and RNA through it glowing was just amazing, had no idea that was something that one could even do
	Genre of APL	descriptive or experimental genre, like an explanation, narrative, argumentative genre, a combination of two genres, descriptive and narrative, descriptive and argumentative, informative, scientific genre, scientifically descriptive, objective and concise, authors state their opinion without bias, narrative with explanations all the way
	Evidence Survey: Positive Student experiences	novelty of the experience cherished, never done something like this before, was fun actually, a new

Part A: From Student Interviews (Continued)

Themes	Findings	Ideas from Participants
		experience, something really different, good, interesting, positive, I learnt a lot, Evidence Survey was not too intense, added to fun and excitement of taking part in this whole activity, electronic access at any time and from any place suitable, hard, it was not too hard or difficult, it was long but important, content knowledge in APL and Concepts of Evidence can be applied in the real world to promote scientific literacy, no pressure to perform and achieve, liked the way everything was spaced out, felt like you are getting something out of it, questions provided food for thought, could do thinking and come back to answer the questions, liked on-line interaction, one was expected to think and not regurgitate facts from memory, liked format of Evidence Survey, reinforced what one was already learning in Biology and Chemistry, video was really helpful in explaining and summariz- the APL, diagrams embedded in the video were illustrative, the pace of survey, the visual and auditory stimuli in combination reinforced the ability to understand and retain what was being explained, came to know a new technology which one could put to use some day, makes one read much more carefully, what one reads makes a lot more sense and therefore one understands better, helpful attitude of all who conducted Evidence Survey, a great opportunity for students, an effort to help students understand concepts which will help

Part A: From Student Interviews (Continued)

Themes	Findings	Ideas from Participants
		them engage better and deeper with Science, a thoughtful gesture promoting the welfare of First Year Undergraduate students taking Science courses, got a peek out of someone's data on a PhD, helped interpret data better, more professional than what I have had to deal with so far, "it gave me something which will always be with me"
	Evidence Survey: Negative Student experiences	rustling noise which could be heard in the video due to the movements of the camera, failing to fully understand the APL though they read it was discouraging, no negative aspects, long, survey was time-consuming, took longer than expected, took away time which could have been spent on own academic studies, unsure of what was expected from the questions, performance on Pre-Test would have been better if it had been preceded by the video (TI), five parts of whole activity should be more closely spaced time-wise and covered within a shorter period of time, overwhelmed with the numerous diagrams, emphasis should be on getting overall picture, technology did not function as well as expected, the time in the semester at which the whole research activity took place
2Efficacy of the Teaching Intervention	Overall impressions of the TI	excellent, very useful, made a big difference, long, very well organized, learnt about self, focussed on presenter, informative, liked it best

Part A: From Student Interviews (Continued)

Themes	Findings	Ideas from Participants
	The TI as a pedagogical tool	really enjoyed it, drawing and flowcharts excellent, prolific representations of MBs binding target, teacher-centred, very useful and informative, using different faculties enhances understanding and retention, like the light came on, excellent, easy to follow, errors explained, pace ideal, can watch when convenient, step-by-step explanation, conversational style, awesome, pointed out what to look for in the APL, definitions of Concepts of Evidence, reinforced understanding and memory at different levels, remedy of great efficacy, it did not hurt, presenting auditorily and visually
	Performance on Post-Test: Student perceptions	performed better due to TI, not done better, do not know about test result, do not know if different, harder questions, a lot easier questions, hoped did better on Post-Test because of TI, probably did better, think so, about the same
3Wrestling with Concepts of Evidence: The views of students	Selecting most important Concept of Evidence	design, validity, Fair Tests and controls (#12), detecting patterns (#19), instruments, calibration and error (#5), design: variable structure (#11), (#12) only a Fair Test can make an experiment in Science valid, setting up a valid experiment is the first step in proper empirical design, if the whole Fair Test is wrong there is nothing one can do about it ... one is done, even if there is a correlation observed in an experiment which is

Part A: From Student Interviews (Continued)

Themes	Findings	Ideas from Participants
		not a Fair Test the relationship will not be seen in reality, you will know that you are actually testing for the right thing, if you have a control you will know that other things are not interfering, if uncontrolled other factors will make you misinterpret your results (#19) detecting patterns and coming up with relationships is most important, if you get a pattern you know that something is going on, shows that one is making progress somewhere, shows some sort of correlation, the more one sees patterns the more one knows that one is heading in the right direction in the study, (#5) if there is a calibration error or some other error, the experiment is not going to work, zeroing is important for accurate and reproducible results (#11) most important factors to consider in design of an experiment are the three different types of variables – Dependent, Independent and Controlled, important to have controls, important to keep certain variables constant throughout the experiment, important to change, observe and control certain variables, relationship between variables must be studied, correlations if any must be detected, established and reported, the absence of correlation is in itself a result that can be reported
	A “Fair Test” defined	a minimally accurate definition was provided, a definition that was of

Part A: From Student Interviews (Continued)

Themes	Findings	Ideas from Participants
		medium accuracy was given, the definition was fully accurate
	Classification of Variables	classified Fluorescence Intensity and Absorbance as Dependent (Responding) variables, Time of Irradiation as Independent (Manipulated) variable
	Manipulating Independent Variables	only one Independent variable changed in an experiment at one time, usually one Independent variable altered at one time, only when one Independent variable varied at one time does the experiment become valid or a Fair Test, when a Fair Test only one Independent variable is changed at one time and all other Independent variables held constant
	Varying two Independent Variables simultaneously	not possible, possible but more difficult, this is an error which undermines the validity of the experiment, if two Independent variables altered at one time one cannot unequivocally assign the change observed to either one of the two Independent variables, as one cannot distinguish between the effects of the two IVs the results are rendered inconclusive and invalid, because you change the Independent variables - like Time of Irradiation and magnesium chloride concentration - one at a time, if you wanted to study the effect of both Independent variables you would have to do two separate experiments,

Part A: From Student Interviews (Continued)

Themes	Findings	Ideas from Participants
4Challenges for attaining substantive knowledge	Dealing with outliers or anomalous data	outliers, errors, erroneous data, abnormal, extra points, interpolated points, a mistake, sources of error repeated the experiment, explained, recognize as weird, redo, call it error, extend the points, go to a logarithmic function, points brought closer to the curve, overlooking could be ignoring potentially ground-changing results, state got obscure results
	Integrity in empirical science	one stray case ignore, anomalies to be noted and experiment repeated, could not ignore outliers, justify or explain them, accountable, results must be real and reproducible
	Identifying "sampling" in experiments	yes, no, so many points, so much time in between, dots seem to be all around certain lines along the x-axis, take a whole bunch of little samples, close to two dozen circles, they have so much to do in a single sample over a period of time, there are diff--erent values for each individual time, no sampling makes no sense, then let the experiment continue on, needed to take smaller samples of original mixture, not taking the whole thing, after each specified period of time you are taking another sample and measuring the FI and Absorbance, don't think you can tell, they were dealing with just one sample, I would say there was sampling, the readings were taken at different times, they are measuring samples here and those are the dots, it is a new sample he takes every time, in

Part A: From Student Interviews (Continued)

Themes	Findings	Ideas from Participants
		the entire figure there are so many points, there would have to be aliquots in order for you to get reliable results out of it, the individual dots are individual samples being taken, the points each of those is a different sample that you are testing, each of these is from the same vial, now I don’t know, impracticality of taking out the whole hybridization mixture, sample being taken out of the original, too time-consuming otherwise, whole experiment can be done at once, can measure the exact time in between the various “scooping out” of the aliquots
	Setting up control experiments	reinforce the data that you have, confirming experiments, supporting experiments, take the magnesium chloride out of the hybridization mixture, with DNA and magnesium chloride, molecular beacons alone, DNA alone, background, control as baseline to be subtracted from sample, set up by removing Independent Variable, you would have to bind it to something else
	Purpose of experiments in APL	finding the optimal conditions Comparing the efficacy of the Molecular beacons probes in measuring photodamage in T & U oligonucleotides, ways to observe damage in DNA, detect the damage using fluorescence, the effect of uv-light on molecular beacons, to see if photo light gives a direct reason why things are the way they are, trying

Part A: From Student Interviews (Continued)

Themes	Findings	Ideas from Participants
		out experiments by having a Control, trying wavelengths and all out, trying to see the damage the sun could do on the skin, trying to find the energy level which affects the DNA, using Molecular Beacons to test the fluorescence, looking at photodamage in T & U, the Molecular Beacons were the probes used, to see if MBs can detect damage done to DNA and RNA, to get more evidence for their theory that they can find mutations

Part B: From Instructor Interviews

Themes	Findings	Ideas from Participants
1Challenges in the content knowledge of the APL	Unjustified omissions	the authors of the Primary Literature were only interested in the change in Fluorescence Intensity, an absolute value was not important since this was only a relative scale, that instrument and run-by or stand-by are experiment-dependent, what was required was just a relative Fluorescence Intensity, the authors are looking at a trend and it is more of a general thing, the arbitrary units are just a function of the fact that they are 'ratioing' their signal versus a fender baseline and so it is a ratio Yarasi et al. (2005) knew that Thymine and Uracil were the most photochemically active bases in DNA and RNA respectively, these authors wanted a very well-defined, simple sequence where they could be sure damage would occur and where the damage is predictive, this was the most controlled and predictive experiment, not sure of the answer, these target nucleotides did not occur in normal DNA/RNA sequences, these two oligonucleotides were well-studied and perhaps were cheaper, these nucleotides may be the ones associated with the kind of damage that takes place, these researchers had to start somewhere
	Elusive details	similarity in shape between the heating and cooling curves indicate that the changes occurring are reversible, this is a folding curve, it folds back to the same shape, it behaves the

Part B: From Instructor Interviews (Continued)

Themes	Findings	Ideas from Participants
		same at all temperatures, the thing is that it follows the same pattern, they are superimposable, like it didn't do anything the processes the Molecular Beacons undergo as the temperature is raised or lowered are reversible, there is no hysteresis, the Molecular Beacons refold spontaneously, to show that they get the same structure back, the Molecular Beacons behave the same way whether they are heated up or cooled down, under the experimental conditions they want to make sure that temperature is not a factor ..... whether it is going up or down
2The Instructors' views on the Teaching Inter -vention	Instructors' opinions on the effectiveness of the Teaching Intervention	the visuals and anecdotal explanations most definitely helped to alter students' understanding of the Concepts of Evidence, students do understand the Concepts of Evidence but unsure whether they were able to connections to their own experimental work, the students had grasped the way experiments are done and the importance of tying experiment ation with the theories, no reservations about the effectiveness of the Teaching Intervention, certainly aware that these particular students would welcome a change and happily watch and learn from the video, these students may even prefer the video to the live teaching that they get in class, the Teaching Intervention definitely has a role in improving the understanding that these students have of these Concepts of Evid-

Part B: From Instructor Interviews (Continued)

Themes	Findings	Ideas from Participants
		ence, I think it will make a huge difference, I predict that the Teaching Intervention will make a huge difference, despite the fact that there was only one, the Teaching Intervention will have tremendous efficacy in furthering the students' understanding of both the Substantive Knowledge of the APL and Procedural Knowledge (i.e. Concepts of Evidence embedded in the APL)
3The thinking gets kind of lost in the content	Teaching Chemistry: the role of Concepts of Evidence	to stipulate a consideration of the factors controlling the collection, measurement and handling of data, this is a good thing because you are going to ask students to think..., in Science you may think you know something but when someone asks "Why? How? Prove it" you cannot, there is value in teaching these Concepts of Evidence, first reaction was to think of referring the paper on Concepts of Evidence by Gott to the teacher teaching the sampling course in that institution, felt this article (i.e. the APL) was a good one and would relate to the undergraduate students, one more experienced Instructor was quick to acknowledge that Procedural Knowledge was never addressed in Chemistry classes in their days "not one iota," Concepts of Evidence can be delved into more deeply, an instrumental course with a stipulated handling of lots of data would be a good course to explicitly introduce these concepts in the First Year of an undergraduate degree, Science teachers could perhaps first mention these Concepts of Evidence in high school

Part B: From Instructor Interviews (Continued)

Themes	Findings	Ideas from Participants
		when students deal with various kinds of variables and sampling issues, I am sure we all have that (i.e. Concepts of Evidence) in one form or the other, all instructors have the goal of imparting knowledge to the students and to achieve this one has to have a plan, the way the Concepts of Evidence are laid out is one such plan, these Concepts of Evidence are currently being taught in stages, the Chemistry students are introduced to Procedural Knowledge probably not as Concepts of Evidence but as a normal part of learning Chemistry, the Concepts of Evidence are not taught as a separate part of the curriculum but they get that piecemeal some application of that in Chemistry, in the General Chemistry class they are introduced to Concepts of Evidence and the Instructors build on what the students learnt in high school, Instructors explicitly emphasize observation, acquiring one's own data and graphing it but they are not introduced to some of the higher order non-linear relationships, multiple measurements, error bars, instruments, calibration and error are only introduced in second-year Analytical Chemistry courses, students did not do a lot of designing their own experiments at the undergraduate level until they reached the fourth year, many fourth year Chemistry undergraduates do two terms of investigative research which entails working one-on-one with a professor, or post-doc or graduate student, the undergraduate student who has been thro-

Part B: From Instructor Interviews (Continued)

Themes	Findings	Ideas from Participants
		ugh such a research project has dealt with some of the higher order Concepts of Evidence, the teaching and learning of Concepts of Evidence are relevant from the First-Year to the Fourth-Year of the Chemistry Undergraduate program, Concepts of Evidence are a part of Chemistry and are little pieces of information introduced here and there
	Teaching more effectively	certainly I think that stuff talking about something like the Concepts of Evidence this is a good thing because you are going to ask students to think about..., my first gut instinct would be to say 'no,' I find it more beneficial for them to discover these things on their own, it probably wouldn't hurt, I think this sort of lays out in detail what ought to be done in order to make it successful and I think it is a good lay out for them, it would not be possible to include all the Concepts of Evidence but may be a few could be introduced, positive that including the Concepts of Evidence into the First-Year teaching of Chemistry would have no adverse effects but had reservations about the number of Concepts of Evidence and the time involved, introducing Concepts of Evidence may increase the students' appreciation and understanding of Science but not convinced that the teaching of Chemistry can be made more effective in this way
4Teaching Concepts of	Two schools of thought	chemists acquire the necessary Procedural Knowledge by working on

Part B: From Instructor Interviews (Continued)

Themes	Findings	Ideas from Participants
Evidence: Instructors' Views	about teaching Concepts of Evidence	projects and ideas struggling with them and dealing with complicated issues, a better way of teaching these Concepts of Evidence would be to keep them below the surface and allow the students to explore them for themselves, it is better to teach those concepts in a "soft" way as undertones to your education not as explicit ideas, the second school of thought was that these Concepts of Evidence have to be addressed at some point and they have to be taught specifically as early as possible, it would be difficult to learn all this material on Concepts of Evidence in one year, it is almost like taking baby steps ... before you can run you have to walk, even in our current structure Concepts of Evidence are covered as a normal part of learning Chemistry, the Chemistry student is taught these concepts "piecemeal," teaching about these Concepts of Evidence should enable students to think on their feet ... how it was designed and how it could be designed better or why there are certain things that are done, uncertain as to whether teaching a unit on just Concepts of Evidence alone would be beneficial for the students, a lot of content that is taught is curriculum-driven, the curriculum is so packed that there is no time to include the Concepts of Evidence, bringing Concepts of Evidence into the lab is also problematic because there are large numbers of students leading to the formation of many sections, consequently achieving consistency is very hard in the

Part B: From Instructor Interviews (Continued)

Themes	Findings	Ideas from Participants
		large First-Year classes
	Level to teach the Concepts of Evidence	should at the First-Year Undergraduate level and not in High School, High School teaching should be limited to the basics, higher level thinking skills such as critical thinking and students judging their work and that of others should be left to the postsecondary institutions, the explicit teaching of the Concepts of Evidence should be commenced only in the First-Year Undergraduate level, students leaving High School should have a very solid and fundamentally basic repertoire of knowledge to fall back on when faced with decisions in Chemistry classes or labs, this seems to be lacking from students graduating from High School today
5Improving the Research Activity	Choosing Primary Literature to write an APL	I wanted to choose a peer-reviewed journal article that had some application and meaning for the general student population at the First-Year Undergraduate level, out of the papers the Chemistry Professor gave me I chose the one in the Journal of Photochemistry and Photobiology titled "Molecular Beacon probes of photodamage in Thymine and Uracil oligonucleotides," the main reason for my choosing this paper was that I thought students would find it interesting because it has the practical application of demonstrating how uv-light induced damage can be detected in DNA or RNA using Molecular Beacon probes which emit striking fluorescence intensity when these nucleo-

Part B: From Instructor Interviews (Continued)

Themes	Findings	Ideas from Participants
		tides are damaged or mutated, I wanted to choose an article which could be understood by First-Year Undergraduate students and I felt this paper fulfilled that condition, the topic of the Primary Literature should be appealing to the students and must catch their attention just like an attractive advertisement, say that paper was on Music or your paper is on Viagra suppose you throw something out like that you can bring in a lot more people, since Chemistry 101is a basic Chemistry course the students may be looking for a paper whose content is related to topics covered in class, the majority of students enjoyed reading the APL and found that its most interesting parts were the meanings and annotations provided in the right column, a minority of students found the initial reading of the APL a struggle and somewhat frustrating, the experiences of this minority will be taken into consideration in the future when choosing a Primary Literature article to write an APL
	Improving the response rate	tying this whole research activity to marks might be a way to increase the response rate, since the students are very mark-and-time-oriented I would say 3% of the grade is probably worth like a $100.00 for them or more than that, select a paper that caters more to what the First-Year Undergraduate students know as Chemistry, one had to choose a Chemistry paper in which all the main Concepts of Evidence were

Part B: From Instructor Interviews (Continued)

Themes	Findings	Ideas from Participants
		embedded, any paper that monitored bodily functions or malfunctions such as diabetes would appeal to the students, one has to really choose an article that sells and appeals to the students, it can be difficult to balance the desire to appeal to the students with the duty to teach the Concepts of Evidence, having the Pre-Test and Post-Test done in class instead of on-line might encourage more students to participate in and complete the tasks
	Collection of evidence in the primary literature	range of temperature, unsure of the answer, because most organisms live between 25^0-60^0C, it is more representative of life as we know it, it could be this "thing" decomposes at about 60^0C, this could be the temperature range of interest for what these authors are studying, could be due to the limitations of the instruments used for measurement, the degradation or denaturation of the Molecular Beacons occurred at temperatures outside the range studied, the room temperature is usually between 20^0-30^0C and that is why Yarasi et al. started measurement at that temperature and since the Fluorescence Intensity is already low at that temperature they did not lower the temperature any further, the Fluorescence Intensity peaked at around 50^0C and then declined and therefore the upper limit was set at 60^0C
	Representation of evidence in the APL	error bars, no proof that these experiments were repeated multiple times, if error was large would be

Part B: From Instructor Interviews (Continued)

Themes	Findings	Ideas from Participants
		difficult to put them on one plot, error bars were not included because: they are smaller than the symbols representing a point on the graph; they are used for fitting purposes to decide which points are real and which have to be modified to fit the curve or line obtained in the graph; error bars are not significant and can be neglected because this is not a quantitative study; here the authors are trying to show a trend rather than absolutely fitting the data points into a line or curve
	Reliability and validity in the primary literature	level of reliability or validity in the paper is sufficient, this is preliminary work and the results obtained since then have been consistent especially the structure of Molecular Beacons and how they work, since this work has been published in a peer-reviewed journal one would assume that other people in the field would deem it sufficiently reliable and valid, some Instructors were unsure of the answer, the heating and the cooling are a means of triangulation – two ways of testing the same thing, is the Fluorescence Intensity really measuring the damage, there is no evidence in the diagrams for reliability of data because it is not clear whether the points shown in the figures are aggregates or single points, the Fluorescence Intensity was a little worrisome and was not sure whether the authors were just conforming to the standards in the area of Photo-Chemistry and Photobiology, the data did not show Error Bars, the

Part B: From Instructor Interviews (Continued)

Themes	Findings	Ideas from Participants
		way the data were presented in the figures gave the impression that the results shown were from just one experiment when that might not have been the case, the authors should have indicated very clearly in the paper how many trials had been done, perhaps the trials had been done a number of times, one wonders how scientific it is when it assigns arbitrary units to the Fluorescence Intensity, they could have put units in and placed a Δ in the end, it is a little bit "iffy," but ...

Appendix J

A Detailed Example of a Conceptual Change

The example cited here is from the interviews with Students 51, 12 and 24 specifically the section on Sampling. As mentioned above, four conditions have to be met before a conceptual change can occur, and each condition is explored in detail below.

Dissatisfaction with existing conceptions.

It is unlikely that a new conception will displace an old one unless the old conception meets with problems along the way and "a new, intelligible and initially plausible conception" is accessible that resolves these problematic issues (Posner et al., 1982, p. 220). In other words, the individual must encounter some dissatisfaction with the conception that exists before this person will give serious consideration to a new conception.

When I first asked Student 12 (S12) whether there was sampling in the experimental set up depicted in Figure 6, the reply given appeared to indicate that Student 12 was initially of the opinion that there was no sampling in Figure 6. This student seemed to believe only one sample was used and it was irradiated for a period of time in the Irradiator, then the whole sample was taken out and measured for Fluorescence Intensity and Absorbance. The same sample was then returned to the Irradiator and the process was repeated for the second period of time. It was also evident that Student 12 was unsure of whether the answer provided was correct. Student 24, on the other hand, when asked whether there was sampling in Figure 6, immediately admitted to not knowing what sampling meant although this student should have known. This demonstrated an inner dissatisfaction with the current knowledge or with existing conceptions.

The initial question I asked Student 51 (S51) was, "Do you think there is Sampling in this experiment? And could you say why?" Student 51 firmly gave the answer:

> I'm going to say "No," because there are different concentrations of the Absorbance and the Fluorescence Intensity. There are 400 nM rU_{17}. That is the hybridization mixture. (p. 26)

Student 51 appears to have the idea that the concentrations of the solutions used to measure Absorbance and Fluorescence Intensity should be the same, and if these concentrations are not identical then Sampling cannot be used. The hindrance to employing Sampling, according to Student 51, seems to be the difference in concentrations between the solutions used to measure Absorbance and Fluorescence Intensity. If one looks at this objectively though, it is clear from the footnote in Figure 6 in the APL (Appendix A) that both Absorbance and Fluorescence Intensity are determined using the same solution and, hence, the concentrations used for both determinations are identical.

Dissatisfaction occurs due to the presence of an anomaly, which arises when "one simply cannot make sense of something" (Posner et al., 1982, p. 220). If one is unable to assimilate something that is thought of as being easily assimilable, then an anomaly exists in that situation. When students cannot explain an anomaly, they experience what is referred to in the literature as "cognitive disequilibrium" or "cognitive conflict" (Piaget, 1985). It is this "cognitive disequilibrium" that provides the impetus for them to learn about or create new concepts (Chinn & Malhotra, 2002, p. 327).

In the case of Student 12, I had to create this "cognitive disequilibrium" by asking:

> EV: ... What do you think? How did they get so many points, and they are all different times, right?
> S12: Oh well, I guess you could tell that time ... I thought, I interpreted it as they were taking one sample and irradiating at different times. (p. 47)

Student 12 had no explanation for why the authors (Yarasi et al., 2005) got so many points in Figure 6. As shown above, Student 12 admitted to having been mistaken and had thought there was only one sample, which was the whole of the Hybridization Mixture. This erroneous assumption, when recognized, created a “cognitive disequilibrium” in Student 12 as shown below:

> Oh well, I guess I could tell that time ... I thought, I interpreted it as they were taking one sample and irradiating at different times. (p. 47)

Student 12 had thought that the same Hybridization Mixture had been irradiated each time for the different periods of time as specified in Figure 6.

Student 51 had evidently reached this state of “cognitive disequilibrium,” when I asked the second time:

EV: So my question is, “Is there Sampling in this experiment?” What do you say?
S51: My first answer was “No,” but now I don’t know ... (p. 26)

This disequilibrium has apparently caused so much confusion that this student does not know what the right answer is any more. It should be stressed here that dissatisfaction/disequilibrium and conflict have the same role in cognitive development. “In both cases they motivate searching; without them, knowledge would remain static” (Piaget, 1985, p. 11).

Subsequently, I asked for the reason Student 51 had said there was no Sampling in this experiment. From the exchange that followed, Student 51 seemed to be preoccupied by this notion of concentrations dictating whether Sampling exists or not, but seemed to be unsure which solutions’ concentrations to consider. At first, Student 51 voiced a real concern that there could be no Sampling because the concentrations were different. However, Student 51 did not specify further which concentrations this student was referring to, and was worried about. When I specifically asked “Which concentrations?” the first response was “the concentrations of the RNA and the Molecular Beacons.” Then, hesitantly, Student 51 modified that response to “RNA and the hybridization mixture.” On further questioning I asked this student where (i.e., between what two solutions) the concentration differences were most crucial. Student 51's final choice was between "the buffer and the hybridization mixture" (p. 26).

Intelligibility of a new conception.

When a student is dissatisfied with existing theories, and attempts to consider an alternative idea, the new notion must seem intelligible to this student. Just because the student finds the alternate conception intelligible, however, does not imply this is sufficient for accommodation to occur (Posner et al., 1982). "Intelligibility also requires constructing or identifying a coherent representation of what a passage or theory is saying" (p. 216). Student 12 arrived at this coherent representation in stages. This student first noticed that the dots (showing the points along the time of irradiation, at which the Absorbance and Fluorescence Intensity were measured) were in very close proximity "almost lying on top of each other" (p. 47). When I then asked this student whether these dots depicted the same Hybridization Mixture being taken out and then put back in, or whether they were different samples (i.e. aliquots) taken from the Hybridization Mixture (i.e. "pot") Student 12 admitted that they must be different aliquots taken from the same "pot." Student 12 identified the factor which was misleading this student's line of thinking: she was misled by the "curve."

In this example, when a small fraction of known volume (i.e., an aliquot) of the hybridization mixture (of known concentration, containing the DNA/RNA in solution), is withdrawn and placed in a test solution (of known volume), how does one determine the concentration of the RNA/DNA in the test solution? This was the anomaly or problem that Student 51 was faced with, and which this student could not understand or explain appropriately. When Sampling, the concentration of the hybridization mixture from which the aliquot was taken and the concentration of the solution into which the aliquot was delivered (i.e., the test solution) were not required to be identical. However, both need to be known (or be readily calculable). I tried to convince Student 51 of this by explaining:

> EV: ... But the hybridization mixture, that is this mixture in here, in the Irradiator, and it has got this concentration and this concentration (pointing to the hybridization mixture in the Irradiator). What is wrong with that? This is at a certain concentration but because of this there is only a small aliquot of this, this is not the whole of the sample, and we know that. So we take an aliquot and we measure the Fluorescence Intensity and the Absorbance of that. (p. 26)

I tried to assure this student that there was nothing wrong with taking this approach of working with two different concentrations in the hybridization mixture and the test solution, as long as we know what they are. However, Student 51 was focussed on the anomaly in the difference in concentrations between the hybridization mixture and the aliquot taken out to make the test solution. Student 51 interrupted my assurance and said," But the concentration will probably be different from the ..." (p. 26), and did not complete the sentence. The pre-instructional conceptions possessed by Student 51, confused and pulled this student back. In an attempt to help Student 51 overcome this confusion, I went on to further explain that the issue being investigated in Figure 6 of the APL was not concentrations, but the effect of increasing the Time of Irradiation on physical properties, such as Fluorescence Intensity and Absorbance of the RNA bound to Molecular Beacons. Numerous studies (Driver, Squires, Rushworth, & Wood-Robinson, 1994; Confrey, 1990; Driver, Guesne, & Tiberghien, 1985) have shown that such pre-instructional conceptions are very resistant to change and keep resurfacing in students' thinking. For example, after studying the topic of photosynthesis, many students continue to hold on to their pre-instructional belief that plants get their food from the soil (Anderson & Smith, 1984).

Initial plausibility of a new conception.

When for the second time I asked Student 24whether there was sampling in Figure 6, this student replied in the affirmative, and reasoned: "I think just through the way the time scale works, there are obviously scoops done at different times, and then taken up for examination." Student 24 justified the plausibility of this new conception by the time scale that was used. The time scale dictates that there would have been different scoops done at the various closely-fit times. The different scoops referred to, implied sampling.

Despite the fact that I had explained that it was perfectly acceptable scientifically to measure a physical property such as Fluorescence Intensity or Absorbance across solutions differing in concentrations, and then correlating them using the dilution factor, Student 51 found that difficult to accept. Regardless of how intelligible I had made the theory behind the applicability of this dilution factor, Student 51found it counterintuitive and was inclined to fall back on the pre-instructional conceptions that this student possessed, which were very resistant to change.

I digressed again to explain how plausible this new conception is by repeatedly reassuring Student 51 that:

> We are not so worried about whether this and this [pointing to the diagrammatic representation of the hybridization mixture and the test solution respectively] are the same. We are taking a fraction of that [i.e., of the hybridization mixture], so naturally the concentration will be less. But we are taking that into account. We know how much volume we took, and we know what the concentration of that is, so we can always work backwards and find ... the concentration of the original [i.e., the test solution]. (p. 27)

I went on to emphasize that the question I originally asked was simply, "Is this kind of Sampling, ... taking fractions out of the original [i.e., hybridization mixture], and measuring the Fluorescence Intensity and Absorbance, is it taking place here?" (p. 27). Student 51 answered in the affirmative which was the complete opposite of the initial answer that there was no Sampling in Figure 6 of the APL. When I asked why, Student 51 said:

> Because you have as the time goes on ... it is not ... it is completely different from when you have to leave it in, then take it out, and then hold it in for a while longer. (p. 27)

The above answer highlights the need for the Hybridization Mixture to be continually in the Irradiator if the irradiation times are to be kept accurate and the required times are to be attained efficiently. Then, at the times indicated in Figure 6 of the APL, aliquots of a measured volume have to be taken out from the Hybridization Mixture. These aliquots, when added to Molecular Beacons, become the test solutions whose Fluorescence Intensity and Absorbance are measured. Student 51 has attempted to explain that this method of withdrawing several aliquots from the same Hybridization Mixture as it remained in the Irradiator, could not be substituted by taking the whole mixture (i.e. the whole "pot") out; adding Molecular Beacons to it; turning it into one test solution; measuring its Fluorescence Intensity and Absorbance; and then putting this Hybridization Mixture (i.e. whole "pot") back into the Irradiator to produce the next test solution. This would be methodologically inaccurate, inconsistent and inefficient, rendering it squarely unscientific. Student 51 seems to be hinting at this, although this was not articulated clearly or in sufficient detail. It must be noted here that when compared to these students' initial views, Students 12, 24 and 51 shifted to a diametrically opposite view in the understanding of what Sampling was and under what conditions it could occur. This change was a result of the interactions between these students' pre-instructional conceptions and the anomalies that these students experienced in this context. It is noteworthy that when the anomalies that these students observed were repeatedly addressed, conceptual change occurred. This is an example of a radical conceptual change referred to earlier as *accommodation*.

Possibility of a fruitful research program.

When I finally asked Students 12, 24 and 51 whether there was Sampling in Figure 6 of the APL, the answer was, "Yes." When I asked why, Student 51 substantiated the answer with, "Because the dots follow the curve and then make a line of best fit" (p. 27). Student 51 is alluding to the fact that the points obtained experimentally follow a certain pattern, and one can obtain a curve of best fit where most of the experimental points lie along the curve. This opens up new possibilities for further research to see whether this experiment can be repeated, thereby enhancing the validity of this experiment and its findings.

When a conceptual change does occur, it seems to be a long, slow process. All students including children, “have a tendency to interpret new situations in terms of what they already know, thus reinforcing their prior conceptions” (Driver et al., 1985, p. 193). The only exception to this rule is when a student is “unable to interpret a situation in a coherent way” (p. 198). When this happens, there may be alternative conflicting interpretations, referred to earlier as anomalies, that the student can give, or on the contrary, the student may not be able to offer any meaningful interpretation for the situation at all (Driver et al., 1985). In such contexts students may perceive the construction of a coherent meaning as a necessity, and this may then provide the necessary conditions for conceptual change to occur, as described in detail using excerpts from the interviews of the students discussed above.