

University of Szeged
Faculty of Arts
Doctoral School of Education
Program of Learning and Instruction

**Assessing Namibian Students' Abilities in Scientific
Reasoning, Scientific Inquiry and Inductive Reasoning Skills**

PhD Dissertation
by
Linus Kambeyo

Supervisor: Prof. Dr. Benő Csapó



Szeged, Hungary, 2018

Table of Contents

Acknowledgements	4
Abbreviations	6
List of Tables	8
List of Figures.....	10
CHAPTER 1 INTRODUCTION	11
1.1 Context of the study	11
1.2 Statement of the problem	13
1.3 Thesis organization	14
CHAPTER 2 LITERATURE REVIEW	17
2.1 Science education, inquiry and reasoning skills	17
2.2 Education structure in Namibia.	19
2.2.1 Basic education in Namibia.....	19
2.2.2 TBA and ICT policy in Namibia	21
2.3 Scientific reasoning vs. scientific inquiry.....	23
2.4 Scientific Reasoning	25
2.4.1 Importance of scientific reasoning skills	29
2.4.2 Scientific reasoning in schoolchildren.....	31
2.4.3 Existing assessment instruments of scientific reasoning skills	32
2.5 Scientific inquiry	37
2.5.1 Inquiry skills in science education.....	39
2.5.2 How to develop the scientific inquiry skills	40
2.6 Inductive reasoning.....	42
2.7 Motivation to learn science.....	45
2.8 Technology-based assessments.....	46
2.9 The need for Technology-Based Assessment (TBA)	48
2.10 Socio economic status (SES).....	49
2.11 Summary	50
CHAPTER 3 AIMS OF THE RESEARCH AND STRUCTURE OF THE EMPIRICAL STUDIES	52
3.1 Research aims and structure of the empirical studies	52
3.2 Research questions and hypotheses	55
CHAPTER 4 METHODOLOGY	58
4.1 Introduction	58
4.2 Samples.....	58
4.3 Instruments	61
4.4 Procedures.....	68

4.4.1	Data Collection	68
4.4.2	Data analysis	69
4.5	Summary	70
CHAPTER 5 THE STUDIES		71
5.1	Sub study 1. Online assessment of scientific reasoning, inductive reasoning skills and motivation to learn science among the Grade 5 & 7 students in Oshana Region	71
5.1.1	Introduction	71
5.1.2	Methodology.....	71
5.1.3	Results	74
5.1.4	Discussions	80
5.2	Sub study 2. The possibilities of assessing students' scientific inquiry skills abilities using an online instrument: a small-scale study in the Omusati Region, Namibia ..	83
5.2.1	Introduction	83
5.2.2	Methodology.....	83
5.2.3	Results	85
5.2.4	Discussions	92
5.3	Sub study 3. The relationships and assessment of scientific reasoning, scientific inquiry and inductive reasoning with students' socio-economic status.	94
5.3.1	Introduction	94
5.3.2	Methodology.....	94
5.3.3	Results	97
5.3.4	Discussions	105
5.4	Sub study 4. The relationships between 8th graders' second language readings skills, inductive reasoning, scientific reasoning and socio-economic status in Namibia ..	107
5.4.1	Introduction	107
5.4.2	Methodology.....	108
5.4.3	Results	110
5.4.4	Discussions	115
CHAPTER 6 CONCLUSIONS, RECOMMENDATIONS, IMPLICATIONS, AND LIMITATIONS		117
6.1	Introduction	117
6.2	Sub study 1 online assessment of the 5th and 7th graders' SR, IR and motivation to learn science.....	117
6.3	Sub study 2 online assessment of scientific inquiry skills of the 9th and 11th grade students	118
6.4	Sub study 3 the relationships and assessment of SR, SI and IR with students' SES.	119
6.5	Sub study 4 the relationships between RC, SR, IR and SES.....	121
6.6	Recommendations	121

6.7	Educational implications	123
6.8	Limitations of the studies.....	124
6.9	Summary	124
PUBLICATIONS RELATED TO THE DISSERTATION		126
REFERENCES.....		128
Appendix A: Scientific reasoning test (5th and 7th grades).....		144
Appendix B: Inductive reasoning test		157
Appendix C: Scientific inquiry skills test		185
Appendix D: Lawson Classroom Test of Scientific reasoning (8th, 10th & 12th grades) .		222
Appendix E: English Reading comprehension test (8th grade).....		232

Acknowledgements

Conducting this doctoral study on a full-time basis while very far away from my family members has not been an easy task especially for my first-born daughter that I left when she was only 3 months old. However, the opportunity availed to me by the Hungarian government was just too good to be turned down. I would have never managed to realise this dream without the large networks of support I have been lucky enough to have in both my professional and personal life.

To the University of Namibia management team, Staff Development personnel and to the then HoD of the Department of Mathematics, Science and Sport Education then Dr Helena Miranda, this dream would not have been realized if you were not so proactive throughout the application process and the administration of the Staff Development issues. Once again thank you so much Dr. Miranda.

I would also like to thank my supervisor and mentor, Professor, Dr. Csapó Benő. It has been a privilege to work so closely with you, and your ideas have had a profound influence on my thinking. I really wish that in the future I will be able to continue the great work you are doing. So, my heartfelt gratitude to you Professor Csapó Benő.

To my mom, dad and my siblings, your relentless encouragements means a lot to me. I know you always want the best for me. This journey was challenging but you always knew how to make me feel positive along the way. Thank you!

A special gratitude goes to my colleague and a friend, Attila Pásztor. Thank you for your patience, the sleepless night we spend, you trying to teach me and explaining about some important facts about quantitative data analysis, and yeah, we managed to have our papers accepted at EARLI and SAARMSTE conference respectively.

I am very much indebted to the Research groups of the University of Szeged, Institute of Education (MTA group). Special thanks and appreciations go to Korom Erzsébet, Nemeth B. Maria, Pásztor Attila, Orosz Gabor and all the background staff and personnel of this wonderful research group. A word of appreciation and gratitude goes to Molnár Edit Katalin for always availing her time to me, despite her busy schedule.

To my wife, Ms. Reginald N. Kambeyo, you must know that I could not have done this without your support. You were really patient with me, when I had to leave you with our newly born daughter, our Nelago. I know that at times I seemed neglectful but this was for us, for our future. The time you brought another joy into our life alone, our second born daughter, Tuhafifa.

I really wished I was there but you sailed through anyways. I couldn't thank you enough for that. Thank you for being the strong pillar that holds our family together.

To the first the first seven Stipendium Hungaricum group, thank you guys for making my stay here so worthwhile. We cried, laughed and of course persevered together. Wishing you success in your future endeavors. To my colleagues, my class mates (2015-2018) you guys were amazing, you were there for me when I most needed it and you have grown to be like a family to me. It was amazing to have people like you on this journey.

Abbreviations

21CS-	21 st Century Skills
AAAS-	American Association for the Advancement of Science
ANOVA-	Analysis of Variance
CBT-	Computer Based Testing
CoV-	Control of Variables
eDia-	Electronics Diagnostic Assessments
GALT-	Group Assessment of Logical Thinking
ICT-	Information Communication Technology
IR-	Inductive Reasoning
IRT-	Items Response Theory
KSAs-	Knowledge, Skills and Abilities
LCTSR-	Lawson Classroom Test of Scientific Reasoning
LTSR-	Lawson Test of Scientific Reasoning
MoE-	Ministry of Education (Namibia)
MTA-	Magyar Tudomanyos Akademia
NGSS-	Next Generation of Science Standards
NIED-	National Institute for Educational Development (Namibia)
NRC-	National Research Council
NSTA-	National Science Teachers Association
OECD-	Organization for Economic Co-operation and Development
PIRLS-	Progress in International Reading Literacy Study
PISA-	Programme for International Student Assessment
RC-	Reading Comprehension
SDDS-	Scientific Discovery as Dual Search
SEM-	Structural Equation Modelling
SES-	Social Economic Status
SI-	Scientific Inquiry skills

SMQ-	Science Motivation Questionnaire
SPS-	Science Process Skills
SR-	Scientific Reasoning
STEM-	Science, Technology, Engineering and Mathematics
SZTE-	Szegedi Tudományos Egyetem
TBA-	Technology Based Assessment
TIMSS-	Trends in International Mathematics and Science Study
TOLT-	Test of Logical Thinking
UNAM-	University of Namibia
USA-	United States of America

List of Tables

Table 1.	The Comparison of Lawson's Classroom Test of Formal Reasoning between the 1978 version and the 2000 version.	36
Table 2.	The timeline and the research activities	54
Table 3.	The samples of the study	60
Table 4.	The distribution of the level of education of the students' parents for study 1.	60
Table 5.	The distribution of the level of education of the students' parents for study 2	61
Table 6.	The distribution of the level of education of the students' parents for study 3	61
Table 7.	The distribution of the level of education of the students' parents for study 4	61
Table 8.	Instruments used in the study	68
Table 9.	The distribution of the level of education of the students' parents	72
Table 10.	The SMQII questionnaire that was developed into an online version (Glynn et al., 2011, p. 1165)	73
Table 11.	The reliability indices and descriptive statistics of the test and the questionnaire	75
Table 12.	Mean scores and correlations between subtests and main construct (scientific reasoning).	75
Table 13.	Descriptive statistics and reliability indices of SMQ	76
Table 14.	SMQ correlations with scientific reasoning	77
Table 15.	Reliability indices and descriptive statistics of the inductive reasoning	78
Table 16.	Correlations between inductive reasoning and scientific reasoning	79
Table 17.	Effect of subscales of inductive reasoning on scientific reasoning	80
Table 18.	Characteristics of the parents' educational level	84
Table 19.	Descriptive statistics of the whole scale and subscales.	86
Table 20.	Correlations coefficients between 7 sub scales and the whole test.	87
Table 21.	Linear regression analysis for scientific inquiry skills as dependent variable	87
Table 22.	The grades mean difference.....	89
Table 23.	The genders mean difference.....	90
Table 24.	The samples of the study	95
Table 25.	The distribution of the level of education of the students' parents for sub study 3....	95

Table 26. Whole sample descriptive statistics for LCTSR.....	96
Table 27. Whole sample descriptive statistics for scientific inquiry skills.....	96
Table 28. Whole sample descriptive statistics for Inductive reasoning skills	97
Table 29. Grade 10 and 12 descriptive statistics of LCTSR.....	98
Table 30. Grade 10 and 12 descriptive statistics of scientific inquiry skills	98
Table 31. Grade 10 and 12 descriptive statistics on inductive reasoning	98
Table 32. The paper and pencil English reading comprehension tests.....	108
Table 33. The reliabilities indices of the reading comprehension tests and its subscales	109
Table 34. Students' achievement in scientific reasoning test.....	111
Table 35. Students' achievement in reading comprehension test.....	111
Table 36. Students' achievement in inductive reasoning test.....	111
Table 37. Distribution of parents' level of education (%).....	114

List of Figures

Figure 1	Structure of Basic Education in Namibia (Ministry of Education, 2010)	19
Figure 2	Definition of Inductive reasoning (Klauer, 1996, p. 38).	44
Figure 3	Namibian Regions Source: www.namibiansafari.com/natravelMAP1.htm	59
Figure 4	Sample task for conservation.....	62
Figure 5	Sample task for proportional reasoning.....	62
Figure 6	Sample task for logical operations.....	63
Figure 7	Sample tasks of the SI tests	64
Figure 8	Sample tasks of the LCTSR tests	65
Figure 9	Sample items (Inductive reasoning) used in all sub studies.	66
Figure 10	Person-item maps for the reasoning skills test	76
Figure 11	The 5 th & 7 th grades performance on scientific and inductive reasoning	78
Figure 12	Students' performance on scientific reasoning and inductive reasoning	79
Figure 13	Sample items; Data handling and hypothesis formulation (Nagy et al., 2015)	85
Figure 14	The person item map of the scientific inquiry skills test.....	88
Figure 15	Rasch analysis of the grades performance on each item	89
Figure 16	Rasch analysis of the gender performance on each item.....	90
Figure 17	Effect of mothers' education level on students' performance on items	91
Figure 18	IRT maps for the three domains	101
Figure 19	A model of SR, SI and IR with dimensions explaining each domain.	102
Figure 20	A grade 10 model of SR, SI and IR with dimensions explaining each domain.	103
Figure 21	A grade 12 model of SR, SI and IR with dimensions explaining each domain.	104
Figure 22	A model of RC, dimensions explaining the achievement of IR.	113

CHAPTER 1 INTRODUCTION

1.1 Context of the study

Several reforms have taken place in the Namibian education system since independence in 1990, particularly in curriculum and assessment areas (Iipinge & Likando, 2012). However, none of the reforms provided explicit guidelines on the assessment of the current trends and issues pertaining to the 21st century. While this thesis was being written, the national curriculum was once again being reformed. These many reforms have required new teaching methods and approaches towards teaching and learning.

One of the issues that needs reform is the assessment of reasoning within science amongst others. It is argued that regular assessment monitors students' cognitive development, and ensure that they possess the reasoning skills necessary for them to understand and master the science learning material in a meaningful way on the one hand, and to check if science education stimulates students' cognitive development as much as it can be expected on the other (Csapó, 2012). This idea is echoed by (Adey & Csapó, 2012; Adey & Shayer, 1994; Csapó & Szabo, 2012), who assert that the content-based methods of enhancing cognition by applying science material for stimulating development provide rich resources for identifying reasoning processes which can be relevant in learning science and which can be developed through science education.

The new premises assessing and evaluation in science education and education in general in the 21st century, caught my attention and triggered the interest in this project. I decided to embark on a project to learn more about the assessment of some of the current trends and issues in the 21st century such as scientific reasoning, scientific inquiry skills and general reasoning skills (inductive reasoning). Some current research in science education have indicated that students need to acquire scientific content that should help them enhance their reasoning skills (Han, 2013). Others, such as the Organization for Economic Cooperation and Development (OECD), developed some science assessment framework tools in order to measure the level of students reasoning skills. These tools measure skills that college and business leaders say young people need to develop, such as those related to reasoning and applying knowledge to solve problems in an unfamiliar context (OECD, 2016).

To succeed in today's globally competitive era, students need to master rigorous academic content and develop twenty-first century skills such as critical thinking and problem solving (OECD, 2016). Hence, helping students succeed at the national and global level requires new

strategies, such as enhancing students reasoning skills and assessing it to know if students are keeping up with contents in this ever-changing world.

Furthermore, tests in scientific reasoning, scientific inquiry and inductive reasoning skills can provide valuable information at various levels as alluded to earlier. Teachers will be able to evaluate and reflect on their teaching styles should the results of the test not being satisfactory. Both teachers and children may be motivated if the results of the test are good. In their research Adey and Csapó (2012) argue that once teachers overcome the urge to teach the reasoning skills directly, they (teachers) will find the results of reasoning tests useful to inform them of where children are positioned so that they can (a) map out the long road of cognitive stimulation ahead (b) better judge what type of activities are likely to cause useful cognitive conflict - both for a class as a whole and for individual children. Moreover, a diagnostic assessment program should support the renewal of primary education (Nagy, 2009). According to (Nagy, 2009) this program should have a dual purpose, it assists individual development by providing learner-level feedback and its summarized results can be used to establish various reference norms. It is further explained that, diagnostic assessment as a direct tool of criterion-referenced education is a method of learner-level evaluation by definition (Nagy, 2009), as such, it is reliant on the longitudinal documentation of individual progress.

The effectiveness of basic education is often evaluated through low-stakes educational assessment studies both at a national and an international level (Vainikainen, 2014). In Namibia, these low stakes assessments occur nationally and assess specific subject content at the end of primary phase, junior secondary phase and secondary phase. It is also worth stating that Namibian students do not participate in any international assessment programs such as Programs for International Student Assessment (PISA), Trends in International Mathematics and Science Studies (TIMSS) or The Progress in International Reading Literacy Study (PIRLS), the neighbouring country South Africa at least does take part in these international programs (Reddy, Zuze, Visser, Winnaar, Juan, Prinsloo, & Rogers, 2015). Therefore, the background knowledge and understanding of what level the students from Namibia are at with regard to scientific reasoning and inquiry skills as well as inductive reasoning skills, in international assessments has never been established even at national level. This research tries to shed some lights on these constructs in a Namibian context. Developed countries like Finland provides an example of a system in which the monitoring of the educational outcomes is based entirely on sample-based

assessments which normally do not have any consequences for the participating students at an individual level (Vainikainen, 2014).

Therefore, this chapter introduces this study, which assessed students' abilities in three constructs which feature in various international assessments programs. The three major constructs assessed are, (Scientific reasoning [SR], Scientific inquiry [SI] and Inductive reasoning [IR]). The study was carried out under four sub studies in the northern regions of Namibia. In this chapter, a problem statement is briefly described, as well as descriptions of the thesis organization.

1.2 Statement of the problem

As stated above, large number of studies have highlighted the importance and benefits of assessing SR, SI and IR. A broad range of instruments, including observation protocol, tests, and item banks, are available which can be used to assess different aspects of general cognitive development as well as specific skills such scientific reasoning skills and scientific inquiry skills which learners are expected to master at school in the 21st century. However, the ability level of students scientific reasoning, scientific inquiry and inductive reasoning skills of the Namibian students have not yet been established. Namibia currently has few mechanisms other than the grade 12 end of school year examination for measuring the performance of the system against international benchmarks (MoE, 2007). Few feedback mechanisms are in place to identify areas of pedagogical difficulty and current testing regimes are not well aligned with modern learning and the pedagogical demands of the 21st century. In order to efficiently foster the thinking and reasoning skills in the classroom context and monitor students' development, reliable and valid, easy to use assessment tools have to be available for the teachers to use.

To date many studies have been conducted on assessing students' scientific reasoning skills, general thinking skills (inductive reasoning), and scientific inquiry skills, mostly in developed countries (Bao et al., 2009; Csapó, 1997; Han, 2013; Mayer et al., 2014; Klauer, 1999, 2001; Wenning, 2007). Results from these studies have shown that assessing these skills helps to improve the teaching and learning in schools, improve the education system in general, and help improve the thinking and reasoning skills of the students in particular. Informed by the literature, this study was deemed necessary to act as a stepping-stone for further research in these areas in Namibia. Therefore, the study wishes to ascertain the thinking and reasoning skills of Namibian students at

specific grades. To achieve this, four sub studies were carried out using the cross-sectional research paradigm.

1.3 Thesis organization

This dissertation consists of six chapters. Chapter One introduces the research, outlining its context and its motivation. An overview of the thesis chapters is also provided.

Chapter Two furnishes a review of literature relevant to the study. The literature reviewed focuses on the main three main constructs studied i.e., scientific reasoning, scientific inquiry and inductive reasoning. The general consensus from the literature is that students need to be taught not only STEM content, but also the skills needed for their survival in the 21st century. Countries are encouraged to carry out low stakes assessments at school and county level, in line with international assessments organizations. A brief synopsis of the importance of technology-based assessment is also discussed in this chapter.

Chapter Three provides the research aims and structure of the study. The research questions and hypotheses are also discussed in this chapter. Chapter Four discusses the methodologies employed to frame, plan and carry out this research study. It also describes the study's research paradigm: a cross sectional quantitative approach was deemed appropriate for this study. Research site and sampling issues are discussed in detail. Data gathering techniques are described, as well as how the data were analyzed and validated.

Chapter Five presents the analyzed data from the four complementary studies used for this research. Each study is discussed in detail according to the research questions guiding that specific sub study. In sub study 1, research has indicated that the advantages of technology-based assessment, such as online test administration and automatic calculation of scoring, reduced the time and cost of the testing process. So, we embarked on this study using technology-based assessment. The first sub study examines pupils' performance in the 5th and 7th grades in cognitive tasks such as scientific reasoning and inductive reasoning. The purposes of this study were to explore the possibilities of online assessment, to examine students' knowledge and thinking skills based on log file analyses, and to explore the relationship between reasoning skills in a science context and motivation to learn science in Oshana region, Namibia. Educational equity is then evaluated by adding gender and mother's educational level. The Rasch model was also used in order to examine the students' ability level and the discrimination power of the tests.

Informed by the results of the first sub study, two assumptions were made. The below average performance on the tests could be caused by the students either being too young or students not have acquired the necessary language skills required to handle the tests. The students may also not have the necessary skills to use computers as the test was online. The second study was then carried out with an older sample (i.e. the 9th and 11th grades). The purpose of this study was to 1) explore the possibilities and feasibilities of an online assessment of scientific inquiry skills, 2) examine the psychometric properties of the test and 3) ascertain the ability level of the 9th and 11th graders scientific inquiry skills. A test of scientific inquiry skills was used, students were required to use their cognitive skills to answer questions based on different subscales of scientific inquiry processes.

Based on the results of the second sub study, where the students' performance improved compared to the 5th and 7th graders, a decision was then made to assess the 10th and 12th grades. This was also informed by the fact that these are the grades that write the national exit examination. This study used paper and pencil methods and examined all the three constructs; scientific reasoning skills, scientific inquiry, and general thinking skills such as inductive reasoning skills, needed by the students in the 21st century. The relationship between scientific reasoning (SR), inductive reasoning (IR) and scientific inquiry (SI) were explored. A structural equation model was used to determine if inductive reasoning impacts the SR and SI. One-parameter Rasch analyses was applied to show item difficulty and students' ability level.

The last sub study, sub study 4, focuses on students reading comprehension skills and its impact on the students' performance in IR and SR. In Namibia, English was made official language at independence in 1990, even though people speak different languages at home. The schools 'language policy states that mother tongue should be used from pre-grades to grade 4, then in grade 5 upwards, the language of instructions should be English. This may present challenges for students as they might not have acquired the necessary language skills for learning. With that in mind, the reading comprehension test was deemed necessary given the fact that some of the students performed below average in the tests. The students' socio-economic status was also explored to ascertain whether it affected students' performance. The first assumption was that students might not have an adequate level of comprehension of English as a medium of instructions which triggered this sub study. Therefore, the aim of this study was to examine the effect of reading comprehension on SR and IR achievement.

Finally, Chapter Six concludes the study by providing a summary of the findings of the study, making certain recommendations arising from the study, describing the limitations of the study and suggesting avenues for future research. The references list followed immediately after the conclusion chapter, and the appendices of the instruments used to collect the data are attached.

CHAPTER 2 LITERATURE REVIEW

2.1 Science education, inquiry and reasoning skills

Current thinking about the desired outcomes of science education is rooted strongly in a belief that an understanding of science should be a feature of every young person's education (OECD, 2013). Indeed, in many countries, science is a key element of the school curriculum from kindergarten until the completion of compulsory education. However, the emphasis on the curricula and its frameworks should not rely on producing individuals who will be producers of scientific knowledge, but rather it should be on educating young people to become informed critical consumers of scientific knowledge, a competency that all individuals are expected to need during their lifetimes (OECD, 2013).

Science education is vital as it i) promotes a culture of scientific thinking and inspires citizens to use evidence-based reasoning for decision making, ii) ensures that citizens have the confidence, knowledge and skills to participate actively in an increasingly complex scientific and technological world (Zhou et al., 2016). Further to that (Turiman, Omar, Daud, & Osman, 2012) recommend that to overcome the challenges of the 21st century in science and technology education, students need to be equipped with the 21st century skills to ensure their competitiveness in the globalization era. Tytler (2007), echoed the same sentiment when entering the debate about the role of school science education.

The 21st century skills in science education that are expected to be mastered by students are comprised of four main domains, digital age literacy, inventive thinking (reasoning), effective communication and high productivity (Turiman et al., 2012). In their report, (OECD, 2013) affirm that many of the challenges of the 21st century will require innovative solutions that have a basis in scientific thinking (reasoning skills) and scientific discovery.

Elsewhere, developers of Australia's national science curriculum identify three possible pathways that students need to be prepared for; to make personal decisions on the basis of a scientific view of the world; to become the future research scientists and engineers; and to become analysts and entrepreneurs in the diverse fields of business, technology and economics (National Curriculum Board, 2009).

Although in Namibia, secondary school teachers historically tend to enact a view that they are preparing students for university, international plans, like the Australian School Science

Education plan 2008-2012, (Goodrum & Rennie, 2007) identify the fundamental purpose of school science education as among others, promoting scientific reasoning and scientific literacy. They further extend these views by stating that science not only prepares students for citizenship but “provides a firm basis for more specialized, discipline-based subjects in upper secondary school that lead to science courses at university and prepares students for technical education courses that lead to science-related careers” (Goodrum & Rennie, 2007, p. 70), thus bringing together both sides of the debate. This focus is in line with National Institute of Educational Development, (NIED)’s (2014) views that scientific and technological literacy is the key purposes for science education for all students, not just those destined for careers in science and engineering, while the National Core Curriculum (2012) for Hungary, proposed that scientific literacy should enable individuals to navigate their way through life, rather than focusing on tertiary studies only.

Furthermore, science education has always been considered to be one of the best tools for cultivating students’ minds. Scientific activities such as conducting empirical research, designing and executing experiments, gaining results from observations and building theories are seen as those in need of the most systematic forms of reasoning (Adey & Csapó, 2012). Elementary science education introduces young children to the basic facts about objects, materials, and organisms as well as the activities involved in designing and conducting a scientific investigation (Lazonder & Kemp, 2012). By engaging in these activities, children can start to develop proficiency in the scientific reasoning skills as well as scientific literacy.

In America, the AAAS, (1990) has stated that learning science should be consistent with the nature of scientific inquiry, meaning that it should begin with questions about nature, concentrate on the collection and use of evidence, including the formulation of arguments from evidence, and be situated within the context of history. This goal was further reinforced by the National Research Council’s [NRC’s] *National Science Education Standards* (1996), which emphasized the importance of evidence in the science classroom when they set out the five essential features of inquiry. A common characteristic of the ‘Essential Features’ is their focus upon the role of evidence in scientific investigations. These five essential features of science inquiry are:

1. Learners are engaged by scientifically oriented questions.
2. Learners give priority to evidence, which allows them to develop and evaluate explanations that address scientifically oriented questions.

3. Learners formulate explanations from evidence to address scientifically oriented questions.
4. Learners evaluate their explanations in light of alternative explanations, particularly those reflecting scientific understanding.
5. Learners communicate and justify their proposed explanations.

The focus of the essential features is on what the students should know about the nature of science itself. (Furtak, Hardy, Beinbrech, Shavelson, & Shemwell, 2008) assert that for middle school students, ‘the emphasis should be on scientific explanations with evidence, have logically consistent arguments, and use scientific principles, models and theories’ (p. 6).

2.2 Education structure in Namibia.

2.2.1 Basic education in Namibia

The following section gives a brief description of the education system in Namibia. Basic Education in Namibia is sub-divided into five phases: Pre-Primary; Lower Primary 1st - 4th grades; Upper Primary 5th - 7th grades; Junior Secondary 8th - grades; and Senior Secondary 11th- 12th grades. Formal Basic Education is compulsory for all from Pre-Primary to the 10th grade, after which there are various opportunities: entry into formal Senior Secondary education, vocational education and training, direct entry to employment, or distance learning. Figure 1 gives a graphical representation of the structure.

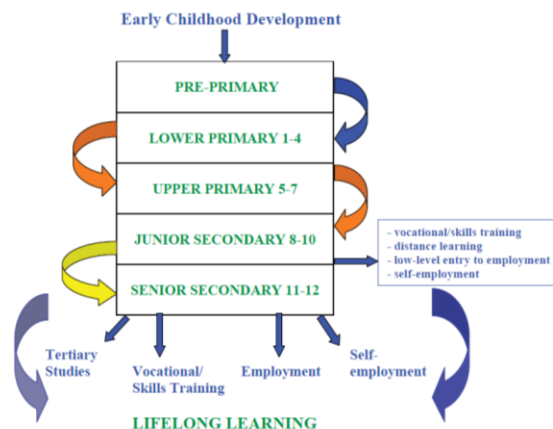


Figure 1 Structure of Basic Education in Namibia (Ministry of Education, 2010)

The Pre-Primary and Lower Primary phases lay the foundation for all further learning. In the Pre-Primary phase, students develop communication, motor and social skills, and concept formation, and are prepared to start formal education. Teaching and learning are through the medium of the Mother Tongue or predominant local language, with a transition to English medium of instructions in the 4th grade (MoE, 2010). The Natural Sciences learning area comprises Environmental Learning (Pre-Primary) and Environmental Studies (1st - 4th grades).

In the Upper Primary phase (5th - 7th grades), students build on this foundation, develop irreversible literacy and numeracy, and develop learning skills and basic knowledge in Natural Sciences, Social Sciences, Technology, Arts, and Physical Education. Teaching is through the medium of English, and the Mother Tongue/predominant local language continues as a subject through to the end of formal basic education. The Natural Sciences learning area in this phase comprises Natural Science and Health Education and Elementary Agriculture (MoE, 2010). As it can be observed, no mention of assessment or training in the 21st century skills are emphasized, the teaching of reasoning skills such as scientific and inductive reasoning is not explicitly enhanced, the focus is mainly on subject content. However, the integration of ICT skills is encouraged although at this stage only in paper, we are yet to see the realisation of this in practical terms.

The Junior Secondary phase (8th - 10th grades) continues with the same learning areas as Upper Primary, consolidates achievements and extends them to a level where the students are prepared for young adulthood and training, employment, or continued formal education. At the end of this phase, those who meet the entry requirements may continue to formal senior secondary education, which provides some specialisation and depth in one area. Those who do not meet the requirements have the option to continue their education through distance education, which enables them to re-enter formal education. The Natural Sciences learning area comprises Life Science, (Agriculture) and Physical Science (MoE, 2010).

In the Senior Secondary phase (11th - 12th grades), all students continue to take English and Mathematics (compulsory), choose a field of study consisting of three mutually supportive subjects, and take supplementary subjects for the 12th grade examinations. At the end of 12th year, students should be well prepared for further study or training, or to enter employment. The Natural Sciences learning area comprises Physical Science, (Agriculture) and Biology (MoE, 2010).

It should be clarified however, that when this dissertation was composed, the basic education in Namibia was being reformed. There would be significant changes to the grouping of the grades.

The implementation of the revised curriculum would be undertaken in phases starting with the lowest grades in 2015 and the final implementation of the revision at the highest grades would be in 2021 (MoE, 2014).

2.2.2 *TBA and ICT policy in Namibia*

Research on technology-based assessments (TBAs) goes back three decades and was originally focused on computer adaptive testing (Almond & Mislevy, 1999; Wainer & Mislevy, 2000; Weiss & Kingsbury, 1984). However, in the late 1990s researchers began investigating how technology could be used to measure complex performances and higher order thinking skills (Baker & Mayer, 1999; Bennett, 1999). In the early 2000s, research was conducted by numerous organizations to explore the potential of TBAs. The U.S. National Assessment of Educational Progress 2001–2003 Technology-Based Assessment Project investigated issues related to measurement, equity, efficiency, and logistical operation in online math and writing assessments (Bennett, Persky, Weiss, & Jenkins, 2007). Findings suggested that although the majority of students reported being familiar with technology, differences in computer proficiency might introduce “irrelevant variance into performance on NAEP mathematics test items presented on computer, particularly on tests containing constructed-response items” (Sandene et al., 2005, p. ix). As computers become more present, familiarity with technology should no longer be an issue; however, poor instructional design, specifically usability and accessibility, can over-load a user’s cognitive resources and impede performance (Sweller, 2005). To this end, the Namibian Ministry of Education formulated an ICT policy for education (2002). The Namibian government through the Ministry of Education committed itself to the provision of ICT infrastructures in schools.

As we move towards a knowledge-based development paradigm, as stipulated in Namibia’s Vision 2030 “Integrating ICT education and training into education and training system”, issues of access to the local and global pool of knowledge and information become paramount (MoE, 2010). The National Information and Communication Technology Policy identifies physical infrastructure and appropriate human capital as the cornerstones for the development and integration of ICT in our society and culture. The Education and Training Sector, long-seen as the torchbearer for capacity development in Namibia, created the ICT Policy for Education to enhance the use and development of ICT in the delivery of education and training.

The five distinct development areas for the use of ICT are:

- Investigation and development of appropriate ICT solutions;

- Deployment of ICT;
- Maintenance and support of ICT;
- ICT literacy; and
- ICT integration.

ICT provides many advantages in the delivery of equitable, quality education, thereby providing an opportunity to improve the lives of our people (MoE, 2010). The need to use new technologies to raise the quality and efficiency of education cannot be over emphasized. It is imperative that we expose our children, parents, and teachers to ICT to improve the quality of education and technical proficiency of our human resources, thus leading to increased productivity and accelerated development. We must also prepare our citizens to adapt to the global economy and participate in electronic commerce.

In addition, we must provide our children with a greater understanding of other peoples and cultures, thus defending our renewed legacy of peace and tolerance (MoE, 2010). The ICT Policy for Education further emphasized that it is intended to coordinate the appropriate development, efficient delivery, and quality use of technology to ensure ICT integration for excellence and equity in education. This policy is an attempt to outline the issue of ICT for Education in the context of the educational sector's struggle to be relevant, responsive, and effective in meeting the challenges of the 21st century. "This policy describes what we want to achieve with ICT in education and what must be in place to achieve it. It does not describe how to achieve it – that is the role of the accompanying Policy Implementation Plan" (MoE, 2010). Researchers elsewhere have also argued that due to innovative possibilities, computer-based assessments offer many advantages over both traditional paper and pencil and the more traditional face to face approach (Pásztor & Molnár, 2013).

Furthermore, the change from the 20th to the 21st century has been accompanied by dramatic changes in virtually all areas of society (Greiff et al., 2014). The globalization and growth of technology have led to fundamental and lasting changes in the societies of the 21st century, also labeled technological societies. Crucially, these changes are reflected in the types of problems encountered in everyday life and thus in demands for the skills students need in order to successfully master life's challenges (Greiff et al., 2014). Whereas factual knowledge is almost instantly accessible nowadays, we are increasingly faced with dynamically changing complex problem environments across a wealth of situations and contexts. It is the mission of education to

adequately supply students not only with factual knowledge and domain-specific problem-solving strategies (which are crucial in and of themselves as well) but also with a broader set of skills required in today's societies (Greiff et al., 2014; MoE, 2010; Molnár, Greiff, & Csapó, 2013).

However, research has revealed many reasons why the use of computer-based assessment is increasing. In the literature, there is mounting empirical evidence that identical paper-based and computer-based tests will not obtain the same results. Such findings are referred to as the “test mode effect” (Clariana & Wallace, 2002). There is no agreed upon theoretical explanation for the test mode effect. Instructional design dogma insists that paper-based versus computer-mediated instructional components should produce exactly equivalent results if the content and cognitive activity of the two are identical (Clark, 1994). In most test mode effect studies, the computer-based and paper-based versions are identical and the cognitive activity should be the same, yet significant differences are regularly observed. For example, paper-based test scores were greater than computer-based test scores for both Mathematics and English tests (Mazzeo, Druesne, Raffeld, Checketts, & Muhlstein, 1991) though other studies have reported no difference between computer and paper-based tests (Schaeffer, Reese, Steffen, McKinley, & Mills, 1993).

Along with the development of information and communication technology (ICT teachers have computer-based testing (CBT) tools at their disposal. However, the same scenario does not exist in the developing world, particularly in Namibia. The popularity of such tools, especially in the sphere of education, stems from the fact that with the use of CBT it is possible to assess more students in less time because the computer program reviews and evaluates tests (Csapó et al., 2015; Pásztor et al., 2015). The advantages of using ICT in education have been recognized in a much broader context than just for the assessment process since the use of ICT makes it possible to replace traditional classroom instructions with e-learning. “Although courses could be realized entirely online by the use of web-based learning environments, a mixed or hybrid approach to eLearning is being used in practice, also called blended learning. This approach combines online learning with traditional face-to-face learning” (Čandrić, Katić, & Dlab, 2014, p. 775).

2.3 Scientific reasoning vs. scientific inquiry

Scientific reasoning is one major component of scientific inquiry, as it contains thinking and reasoning skills (Zimmerman, 2005). This statement by Zimmerman (2005), suggests that there is a thin line between the two constructs. Mayer (2007) describes scientific reasoning with the following processes: formulating scientific questions, generating hypotheses, planning

investigations, analyzing data and making conclusions. Mayer's definition may also fit well with what can be described as scientific inquiry. Klahr (2000) on the other hand, describes scientific reasoning as a process of dual search, which includes searching the 'hypothesis space' and the 'experiment space'. The 'evidence evaluation process' (Klahr 2000) mediates the two steps. Zimmerman (2005) further describes it using similar aspects such as asking questions, hypothesizing, recording and interpreting data. From their definition, three main processes of scientific inquiry emerged. These approaches are central to scientific reasoning: asking questions and formulating hypotheses, planning and performing an investigation, and analyzing data and reflecting on the investigation (Nowak, Nehring, Tiemann, & Upmeyer zu Belzen, 2013). Many different skills of scientific inquiry can be found contained in these three main processes.

Furthermore, scientific inquiry is the way that natural scientists try to answer scientific questions. Scientific inquiry processes can be described as a problem-solving task (Klahr, 2000). As alluded to earlier, it is a circular process, in which questions are asked, investigations are carried out and evidence is evaluated (Mayer, 2007; Zimmerman, 2005). However, other researchers found that scientific inquiry is not a homogeneous construct (Lederman & Lederman, 2012). It consists of a variety of different processes such as methodological and cognitive skills, just like what (Mayer, 2007; Klahr, 2000; Zimmerman, 2000) have alluded to, scientific inquiry is the process that involves research questions and hypothesis generation, planning investigations and identifications of variables, data analysing and generating conclusions in a systematically manner (Lederman & Lederman, 2012). Furthermore, it argued that, to be scientifically literate, one needs to understand how scientists work. Studies have shown that students have difficulties in thinking and working scientifically (Gott & Duggan, 1998; Klahr, 2000; Zimmerman, 2005) because of the systemic way of doing and following procedures indubitably.

Previous studies about scientific inquiry have focused on a specific subject (Mayer, 2012), on one inquiry method (Hammann, Phan, Ehmer, & Grimm, 2008), or on cognitive (Klahr, 2000) or practical aspects of scientific inquiry (Gott & Duggan, 1995). Other studies on scientific inquiry concerned views and perceptions of either students or teachers towards scientific inquiry (Gaigher, Lederman, & Lederman, 2014; Schwartz, Lederman, & Lederman, 2008). Furthermore, many research studies on inquiry skills appeared to be dominated by a focus on classroom-based science investigations (Capps & Crawford, 2013; NRC, 2012). Chinn and Malhotra (2002) found that 'many scientific inquiry tasks given to students in schools do not reflect the core attributes of

authentic scientific reasoning' (p. 176) and suggest that inquiry tasks should go beyond hands-on activities to also include evaluation of evidence, complex data and simulations. Additionally, most of the studies were conducted in European or Asian countries. In this study however, all these various aspects which could influence students' abilities, are considered and the focus is not on a specific subject but on general inquiry and reasoning skills concerning science.

In this study, scientific reasoning is described in terms of the tests taken by the students. The scientific reasoning tasks assess the cognitive and psychological dimensions, while the inquiry skills test assesses the application of science knowledge. In other words, SR tests the cognitive part while the SI tests the practical (experimental part) of science. However, there may be cognitive processes involved in inquiry tests as well, just like in any tests, but the tasks concentrate more on the practicality part of science while the scientific reasoning concentrate more on the cognitive part. Both constructs cognitive (scientific reasoning) and practical skills (inquiry skills methods) of scientific inquiry need to be enhanced, measured and embedded in the science education of the 21st century.

2.4 Scientific Reasoning

Scientific reasoning, can be defined as “formal reasoning” (Piaget, 1970) or “critical thinking” (Hawkins & Pea, 1987), where students are required to apply the domain specific knowledge to domain general. In early studies, SR represents the ability to systematically explore a problem, formulate and test hypotheses, control and manipulate variables, and evaluate experimental outcomes (Bao, Cai, Koenig, Fang, Li, & Wu, 2009; Zimmerman, 2007). Scholars in this field have argued that scientific reasoning and scientific inquiries skills are two sides of the same coins (Han, 2013; Khun, 2007; Zimmerman, 2007). The difference lies in the assessment tasks, SR is more domain general while SI is subject specific. Furthermore, scientific reasoning represents a set of domain general skills involved in science inquiry supporting the experimentation, evidence evaluation, inference and argumentation that lead to the formation and modification of concepts and theories about the natural and social world (Han, 2013; Lawson, 2000).

Scientific reasoning can also be operationally defined as a set of basic reasoning skills that are needed for students to successfully conduct scientific inquiry, which includes exploring a problem, formulating and testing hypotheses, manipulating and isolating variables, and observing

and evaluating the consequences (Lawson, 2000). In terms of assessment, the Lawson's Test of Scientific Reasoning (LTSR) provides a solid starting point for assessing scientific reasoning skills (Lawson, 1978, 2000). The test is designed to examine a small set of dimensions including (1) conservation of matter and volume, (2) proportional reasoning, (3) control of variables, (4) probability reasoning, (5) correlation reasoning, and (6) hypothetical-deductive reasoning. These skills are important concrete components of the broadly defined concept of scientific reasoning ability. Research on scientific reasoning is rooted in the early studies on the cognitive development. Traditionally, the Piagetian clinical interview is used to assess students' formal reasoning abilities. In Piaget's cognitive developmental theory, an individual move to the next cognitive level when presented with challenges in the environment that cause him or her to change, to alter his or her mental structures in order to meet those challenges (Fowler, 1981). Consistent with research, SR tasks are designed to arouse students thinking from their comfort zone of content knowledge and apply it to general reasoning skills. Piaget used the word schema to refer to anything that is generalizable and repeatable in an action (Piaget & Inhelder, 1969). As children grow and mature, these mental structures are described as organized abstract mental operations actively constructed by the children.

As their cognitive structures change, so do their adaptation techniques, and these periods in a child's life are referred to as stages. The first is the sensorimotor stage of the children 2 years of age and younger (Piaget & Inhelder, 1969), an important period of time when the child is constructing all the necessary cognitive substructures for later periods of development. These constructions, without representation or thought, are developed through movement and perceptions. The movements and reflexes of the child in this period form habits that later result in intelligence. This happens through 6 successive sub-stages: modification of reflexes, primary circular reactions, secondary circular reactions, coordination of secondary schemas, tertiary circular reactions, and the invention of new means through mental combinations (Millar, 2002). During this stage, three important concepts are believed to be acquired (a) object permanence, when the child understands the object did not cease to exist just because it is hidden from view; (b) space and time, important to solving "detour" problems; (c) causality, which is when the child begins to realize cause and effect by his or her own actions and in various other objects (Piaget & Inhelder, 1969).

The second is the preoperational stage of 2- to 7-year-old children, transitions from the sensorimotor period with the development of mental representations through semiotic function, where one object stands for another (Millar, 2002). Signs and symbols are learned as similar objects and events that signify real ones. Though mental representation has advanced from its previous stage, children in this period cannot think in reversible terms (Piaget & Inhelder, 1969). Millar helps to describe other characteristics of this level, including rigidity of thought, semi-logical reasoning, and limited social cognition. The rigidity of thought is best described with the example of two identical containers that have equal amounts of liquid. When the contents of a container are poured into a thinner and taller container or shorter and wider container, children at this level freeze their thought on the height and assume the volume is more or less, depending on the height of the container. The height becomes their only focus, rather than the transition of volume. If the liquid is poured from one container into another, children focus on the states of the containers rather than the process of pouring the same amount of liquid.

Cognitively, children are unable to reverse the direction of the poured liquid and imagine it being poured back into the original container which would contain the same amount. They can, however, understand the identity of the liquid, that it may be poured from one container to another and still be the same kind of liquid. In this level, causal relationships are better understood outside of self, as pulling the cord more makes the curtain open more, though they may not be able to explain how it happened. Rather than thinking logically, children in this level reason semi-logically, often explaining natural events by human behavior or as tied to human activities (Millar, 2002).

Most children in ages 8 to 11 are often categorized as in the concrete operational stage in Piaget's theory of cognitive development. According to Millar (2002, p. 52), the mental representations of children in this concrete operational period come alive with the ability to use operations, "an internalized mental action that is part of an organized structure." In the example of the liquid in containers, children now understand the process and can reason the liquid is the same amount though in different sized containers. This ability to use operations may come at different times during this period. Children in this concrete stage begin to better understand reversibility and conservation. Classifications based on the understanding of sizes of an included class to the entire class are achieved (Piaget & Inhelder, 1969). Relations and temporal-spatial representations are additional operations evident in concrete operational children (e.g., children can understand

differences in height and length and include the earth's surface in drawing their perception of things). All of these operations strengthen gradually over time.

The formal operational period is the fourth and final of the periods of cognitive development in Piaget's theory (Piaget & Inhelder, 1969). This stage, which follows the concrete operational stage, commences at around 11 years of age and continues into adulthood. In this stage, individuals move beyond concrete experiences and begin to think abstractly, reason logically and draw conclusions from the information available, as well as apply all these processes to hypothetical situations. Rather than simply acknowledging the results of concrete operations, individuals in this final period can provide hypotheses about their relations based on logic and abstract thoughts. These abstract thoughts look more like scientific method than in previous periods. In the concrete operational period, children could observe operations but lack the ability to explain the process. In the formal operational period, they are able to problem-solve and imagine multiple outcomes. One of Piaget's common tasks in determining if a child has reached formal operational thought is the pendulum problem. The formal operational thinker demonstrates hypothetico-deductive thought by imagining all of the possible rates that the pendulum may oscillate, observing and keeping track of possible results, and ultimately arriving at possible conclusions (Piaget & Inhelder, 1969). As adolescents grow into adulthood and throughout adulthood, formal operations are still developing and abstract thought is applied to more situations. Millar contends Piaget ended his periods of the developmental of logical thought with formal operations. Beyond this point, individuals' thought only changes in content and stability rather than in structure.

In the early works on the measurement of cognitive development, Piaget used multiple problems to test a child's operations of thought (Piaget & Inhelder, 1969). Millar (2002) defined Piaget's methodology as the "clinical method," which involves a chainlike verbal interaction between the experimenter and the child. In this interaction, the experimenter asks a question or poses a problem, and the subsequent questions are then asked based on the response the child gave to the previous question. Piaget developed this interaction in order to understand the reasoning behind the children's answers.

Cook and Cook (2005) noted that through Piagetian tasks, Piaget could better understand preoperational children's thinking. He found these children showed centration, focusing on only

one thing at a time rather than thinking of several aspects. This means they were centered on the static endpoints, the before and after, rather than the process.

The next aspect of logical thinking observed in Piaget's finding was preoperational children's lack of a sense of reversibility. The task of liquid conservation is simple to the logical thinking child. Water from a short and wide container is poured into a tall and skinny container. A preoperational thinker would focus only on the height of the liquid and the fact that the water was first low, then it was at a higher level in the second container; therefore, there must be more water in the second container. With a lack of a grasp for reversibility, the preoperational child does not have true operational thought to allow him or her to imagine the pour reversed and realize the same amount of water is in both containers. The other two conservation tasks are similar to the liquid task. They each show a beginning state, a transformation, and an ending state where something has changed. The importance of children's operational and newer logical thought "is not so much that children are no longer deceived by the problem, but rather that they have now learned some basic logical rules that become evident in much of their thinking" (Lefrancois, 2001, p. 383).

2.4.1 Importance of scientific reasoning skills

The value of science and mathematics education is emphasized worldwide as evidenced in large-scale international studies such as TIMSS and PISA that continually make use of science and mathematics questions. Here it should be noted that Namibia does not take part in these international assessments so one can only guess how the students would perform if developed nations such as USA, UK, and others are ranked low compared to Singapore (PISA, 2015 and TIMSS, 2015). These results have led to demands for the implementation of a more extensive basic education curriculum in science, technology, engineering, and mathematics (STEM). Educational reforms stress the need for an equipped 21st-century workforce, which translates into students learning not only science content but also acquiring advanced transferable reasoning skills (Kuhn, 2010). The development of these skills will better enable students to handle open-ended novel situations and design their own investigations to solve scientific, engineering, and social problems in the real world (Bao et al., 2009).

As science education continued to become fundamental to modern society, there is a growing need to pass on the essential aspects of scientific reasoning and with it the need to better impart such knowledge. Worryingly, the current style of the content-rich STEM education, even

when carried out at a rigorous level, has little impact on the development of students' scientific reasoning abilities (Bao, et al., 2009). The finding from their comparative study (Bao et al., 2009) between American students and Chinese students indicate that it is not what we teach, but rather how we teach, that makes a difference in students' abilities in scientific reasoning. They further determined that students ideally need to develop both content knowledge and transferable reasoning skills (Bao et al., 2009). The onus is upon researchers and educators to invest more time in the development of a balanced method of education, such as incorporating methods that enhance scientific reasoning skills. Previous studies have indicated that scientific reasoning is critical in enabling the successful management of real-world situations in professions beyond the classroom (Han, 2013). For example, in the K-12 education in the United States of America (USA), the development of scientific reasoning skills has been shown to have a long-term impact on students' academic achievement (Adey & Shayer, 1994). Positive correlations between students' scientific reasoning abilities and measures of students' gains in learning science content have been reported (Coletta & Phillips, 2005), and reasoning ability has been shown to be a better predictor of success in Biology courses (Lawson, 2000).

The above findings support the consensus of the science education community on the need for the basic education (Grade 1-12) students to develop an adequate level of scientific reasoning skills along with a solid foundation of content knowledge. Zimmerman (2007) claims that investigation skills and content knowledge bootstrap one another, creating a relationship that underlies the development of scientific thinking. Research has been conducted to determine how these scientific thinking skills can be fostered and which teaching strategies contribute most to learning, retention, and transfer of these skills (Osborne, 2013). For instance, Zimmerman (2007) in her research conducted in Illinois, United States of America (USA), found that children are more capable of scientific thinking than was originally thought, and that adults are less so. She also states that scientific thinking requires a complex set of cognitive skills, the development of which requires much practice and patience. It is therefore important for educators to understand that scientific reasoning ability is best developed through science inquiry-based education at school level. Although there exists a wide range of understandings of what constitute scientific reasoning, the literature seems to generally agree that SR is intertwined with science inquiry. Therefore, a good understanding of the nature of scientific reasoning requires an extended knowledge of science inquiry.

2.4.2 Scientific reasoning in schoolchildren

Traditionally, developmental psychologists have considered the thinking and reasoning of elementary school children as deficient and have argued that scientific reasoning skills emerge only during adolescence (Inhelder & Piaget, 1958). However, in the last 20 years, further developmental research has revealed evidence for early competencies in these skills (Csapó & Szabo, 2012; Mayer et al., 2014). It is further argued that if the required foundations are not constructed, serious difficulties may arise at the later stages of learning, as failures suffered during the first years of schooling will affect children's attitudes towards education for the rest of their lives (Csapó & Szabo, 2012). The development of concepts related to science begins even before the start of formal education and the first years of schooling play a decisive role in steering conceptual development in the right direction. Early science education shapes children's thinking, their approach to the world and their attitudes towards empirical discovery (Csapó & Szabo, 2012).

Moreover, research has also found that even pre-school children understand the relation between covariation data and causal belief when only potential causal factors (e.g., red or green food) covaried partially or perfectly with outcomes (good or bad teeth) (Osborne, 2013). When the effects of more than two variables must be considered, young children often fail to interpret patterns of empirical evidence (Kuhn, 2010). Unlike adolescents or adults, children tend to neglect or distort data, when covariation evidence does not agree with their prior beliefs or knowledge (Molnár, Greiff, & Csapó, 2013). Therefore, research findings indicate that basic experimentation and evidence evaluation skills in pre-school and primary school children do exist (Mayer et al., 2014). The onus is upon us teachers and researchers to develop and assess the scientific reasoning in children at an early stage in their schooling with the view to enhancing learning. When children's scientific reasoning and thinking skills are assessed it would also inform the teachers and parents on the best possible ways to help the children achieve the required reasoning skills.

The goal of general primary education is to develop the basic skills upon which all later learning is built (Nagy, 2009). It is further argued that key areas such as linguistic skills, counting and numeracy and the acquisition of reasoning skills, including deduction, are prerequisites for understanding and mastering scientific knowledge taught at later stages (Nagy, 2009). Nagy (2009), further asserts that often schools offer very few opportunities for children to engage in activities developing the various social roles and social skills needed for cooperation. Research has indicated that the first few years of schooling are decisive with respect to later studies. This is

the stage when the basic skills and abilities which provide the foundations of all later studies are developed (Csapó & Szabo, 2012; Molnár et al., 2013; Nagy, 2009). At this stage, children's overall relationship with learning is shaped and focused, that is to say, learning habits and attitudes towards school and school subjects are formed.

2.4.3 Existing assessment instruments of scientific reasoning skills

What are the possible mechanisms of assessing and testing scientific reasoning then? Guided by Piagetian tasks, a number of researchers (Lawson, 1978; Shayer & Adey, 1994; Tisher & Dale, 1975) have developed their own measurements in assessing students' scientific reasoning abilities, such as the Group Assessment of Logical Thinking Test (GALT) (Roadrangka, Yeany, & Padilla 1982), the Test of Logical Thinking (TOLT) (Tobin & Capie, 1981), and the Lawson's Classroom Test of Scientific Reasoning (LCTSR) (Lawson, 1978). All these tests were administered using paper and pencil methods. For my research study, Lawson test has been adapted and used as it is in the paper and pencil format. Below, I will briefly review the three instruments and their measures.

2.4.3.1 Group Assessment of Logical Thinking (GALT)

Roadrangka et al. (1983) compiled reliable and valid test items for the Group Assessment of Logical Thinking (GALT). In the pilot testing, Piagetian interview tasks were administered to a sub-sample of students for purposes of validation. The 21- item GALT test consist of the first 18 items about multiple-choice problems to be answered by students as well as a selection of reasoning choices to support his or her answer. The final three items are scored upon the child's inclusion of all possible answers and patterns to classify these answers. GALT measures 6 logical operations, including conservation, correlational reasoning, proportional reasoning, controlling variables, probabilistic reasoning, and combinatorial reasoning. They also used a multiple-choice style to present answers and possible reasoning behind those answers. The GALT is sufficiently reliable and valid in its ability to distinguish between students at Piagetian stages of development. Reliability was tested by administering the GALT to students and administering Piagetian Interview Tasks to a sub-sample of those students. They found a strong correlation, ($r = .80$) (Roadrangka et al., 1983). The question selection derived from other reliable and valid instruments helped make this a reliable and valid assessment. The Cronbach's reliability coefficient for internal

consistency of the GALT was reported to be around (.62-.70 Cronbach alphas) (Bunce & Hutchinson, 1993).

One of the six modes measure concrete operations and the other five measure formal operations (Bunce & Hutchinson, 1993). The answers to the GALT items 1 to 18 were considered correct only if the best answer and reason were both correct. For item 19, children must (1) show a pattern and (2) have no more than one error or omission, and for item 20, children must also show a pattern in answers given, having no more than two errors or omissions. To be labeled as concrete operational thinkers, the children had to score 0 to 4.

Transitional thinkers were indicative of the score 5 to 7, and abstract operational thinkers were those children who scored 8 to 12 (Roadrangka et al., 1983). Researchers, predominantly in the field of science education have utilized the GALT to determine a developmental level to gauge student performance, phases in the learning cycle, and cognitive/motivational characteristics. In addition, researchers have administered the GALT to determine the best method of teaching a particular subject based on the students' logical thinking ability (Niaz & Robinson, 1992; Allard & Barman, 1994; Kang, Scharmann, Noh, & Koh, 2005). Through the use of the GALT test, Allard and Barman assessed the reasoning of 48 college biology students and found 54% of these students would benefit from concrete methods of instruction. Sampling 101 more science students in a basic science course showed these researchers that 72% of these students would benefit from concrete methods rather than a traditional lecture approach in the classroom.

2.4.3.2 The Test of Logical Thinking (TOLT)

The Test of Logical Thinking (TOLT) is a 10-item test developed by Tobin and Capie (1981). It measures five skill dimensions of reasoning including proportional reasoning, controlling variables, probabilistic reasoning, correlational reasoning, and combinational reasoning. A high internal consistency reliability ($\alpha = 0.85$) and a reasonably strong one-factor solution obtained from factor analysis of performance on the 10 items suggested that the items were measuring a common underlying dimension. The TOLT test items bare many similarities to the ones used in GALT and Lawson's test.

2.4.3.3 Lawson's Classroom Test of Scientific Reasoning (LCTSR)

Lawson (1978) originally designed his test of formal reasoning to address the need for a reliable, convenient assessment tool that would allow for diagnosis of a student's developmental level. A valid form of measurement prior to the Lawson Test was the administration of Piagetian tasks. This method, however, is time-consuming and requires experienced interviewers, special materials and equipment. A paper and pencil test would be more practical for classroom use, but there are also problems with this method. Paper and pencil tests require reading and writing ability, test takers have no added motivation from the use of materials or equipment, and it is not as relaxed as a clinical interview setting (Lawson, 2000).

In the development of his test, Lawson (1978) aimed for a balance between the convenience of paper and pencil tests and the positive factors of interview tasks. He studied eighth- through tenth-grade students to determine their scientific reasoning skill level. Lawson breaks scientific reasoning into several categories: isolation and control of variables, combinatorial reasoning, correlational reasoning, probabilistic reasoning, and proportional reasoning. Test items were based on these dimensions. The original format of the test had an instructor perform a demonstration in front of a class, after which the instructor would pose a question to the entire class and the students would mark their answers in a test booklet. The booklet contained the questions followed by several answer choices. For each of the 15 test items, students had to choose the correct answer and provide a reasonable explanation in order to receive credit for that item.

To establish the validity of his test, Lawson (1978) compared test scores to responses to interview tasks, which were known to reflect the three established levels of reasoning (concrete, transitional, formal-level). He found that the majority of students were classified at the same level by both the test and interview tasks but that the classroom test slightly underestimated student abilities. Validity was further established by referencing previous research on what the test items were supposed to measure as well as performing item analysis and principal-components analysis. Researchers who used this test have evaluated the reliability of Lawson's test (Version, 2000). Typical internal consistency in terms of Cronbach's α range from 0.61 to 0.78 (Lee & She, 2010).

The popular version of Lawson's Classroom Test of Scientific Reasoning was released in the year 2000. This is also the test used in this study. It is a 24-item two-tier, multiple-choice test. Treagust (1995) describes a two-tier item as a question with some possible answers followed by a second question giving possible reasons for the response to the first question. The reasoning

options are based on student misconceptions that are discovered via free response tests, interviews, and the literature. In the 2000 version, the combinational reasoning is replaced with correlation reasoning and hypothetic-deductive reasoning. The test is also converted into pure multiple-choice format containing 24 items in 12 pairs, (see Table 1). With a typical two-tier structure, the first 10 pairs (items 1-20) each begin with a question for a reasoning outcome followed by a question soliciting students' judgment on several statements of reasoning explanations. Items 21-24 are also structured in two pairs, designed to assess students' hypothetical-deductive reasoning skills concerning unobservable entities (Lawson, 2000). Partially due to the pathways of hypothesis testing processes, these two pairs follow different response patterns. In the item pair of 21-22, the lead question asks for the selection of an experimental design suitable for testing a set of given hypotheses. The follow-up question asks students to identify the data pattern that would help draw a conclusion about the hypotheses. In the item pair of 23-24, both questions ask students to identify the data pattern that would support the conclusions about the given hypotheses. The Lawson's test is widely used in the science education community, hence the need to test it in Namibia in order to determine the scientific reasoning abilities of students.

Table 1. The Comparison of Lawson's Classroom Test of Formal Reasoning between the 1978 version and the 2000 version.

Scheme tested	Item number (1978)	Item number (2000)	Nature of task
Conservation of weight	1	1, 2	Varying the shapes of two identical balls of clay placed on opposite ends of a balance.
Conservation of volume	2	2, 3	Examining the displacement volumes of two cylinders of different densities.
Proportional reasoning	3, 4	5, 6, 7, 8	Pouring water between wide and narrow cylinders and predicting levels.
Proportional reasoning	5, 6		Moving weights on a beam balance and predicting equilibrium positions.
Control of variables	7	9, 10	Designing experiments to test the influence of length of string on the period of a pendulum
Control of variables	8		Designing experiments to test the influence of weight of bob on the period of a pendulum
Control of variables	9, 10		Using a ramp and three metal spheres to examine the influences of sphere weight and release position on collisions.
Control of variables		11, 12, 13, 14	Using fruit flies and tubes to examine the influences of red/blue light and gravity on flies' responses.
Combinational reasoning	11		Computing combinations of four switches that will turn on the light.
Combinational reasoning	12		Listing all possible linear arrangements of four objects representing stores in a shopping center.
Probability	13, 14, 15	15, 16, 17, 18	Predicting chances of withdrawing colored wooden blocks from a sack.
Correlational reasoning		19, 20	Predicting whether correlation exists between the size of the mice and the color of their tails through presented data.
Hypothetic-deductive reasoning		21, 22	Designing experiments to find out why the water rush up into the glass after the candle goes out.
Hypothetic-deductive reasoning		23, 24	Designing experiments to find out why the red blood cells become smaller after adding a few drops of salt water.

Furthermore, (Mayer et al., 2014) also suggested a variety of task formats that can be used to explore scientific reasoning competencies in young children. Apart from self-directed experimentation tasks in which participants may be involved in hands-on physical activities, tasks using story problems are common measures of scientific reasoning. Additionally, contextual support (abstract vs. concrete), task complexity (single-variable vs. multivariable), plausibility of factors, response format (choice vs. production), strength of prior belief or prior content knowledge in scientific domains (e.g., physics, chemistry and biology) have been shown to influence performance in scientific reasoning tasks (Lazonder & Kemp, 2012; Adey & Csapó, 2012). Predict-Observe-Explain (POE) items ask children to make informed predictions about a presented situation (Fu, Raizen, & Shavelson, 2009), and following an observation or summary of what

happens and ask students to provide explanations. For example, students might be asked to predict whether a given object would sink or float in water. Once they found out that the object did sink or float, they must explain why this occurred. This provides opportunities to reliably capture how students reason through and justify their predictions and explanations (Fu et al., 2009).

2.5 Scientific inquiry

Another construct deemed important in science education is that of inquiry skills. Scientific inquiry as a component of scientific literacy has been variously defined. For instance, the National Science Education Standards (NRC, 1996) defines scientific inquiry as follows, “Scientific inquiry refers to the diverse ways in which scientists study the natural world and propose explanations based on the evidence derived from their work. Inquiry also refers to the activities of students in which they develop knowledge and understanding of scientific ideas, as well as an understanding of how scientists study the natural world” (p. 23). It must be noted that literature have asserted that scientific reasoning and scientific inquiry are terms that are intertwined. The slight difference is that, the inquiry tasks are science subject bound (physics, chemistry and biology) while the scientific reasoning tasks are domain general (more cognitive in nature) (Wenning, 2007).

Scholars assert that the use of the term inquiry in science education dates back to the middle of the nineteenth century. Thereafter the term became central to reforms in science education and its meaning was broadened to accommodate various perspectives (Bybee, 2000; De Boer, 2004). Anderson (2007, p. 808) describes inquiry as both a ‘catch phrase’ and a ‘useful label’ to ‘integrate many facets of educational practice’. In a broad sense, scientific inquiry represents the systematic processes of investigating questions about the natural world, leading to the discovery and establishment of new scientific knowledge. In school curricula, scientific inquiry is essential to the development of future generations of scientists, as well as to the development of a scientific-literate population (Antink, & Bartos, 2012; Driver, Leach, Millar, & Scott, 1996; Lederman, Millar, 2006; Millar & Osborne, 1998).

Scientific inquiry has always been an integral part of scientific literacy (Bybee, 2009). Hence, scientific inquiry has been a long-standing area of research and discussion in science education (Fenichel & Schweingruber, 2010; Yeh, Jen, & Hsu, 2012). Scientific inquiry is seen as a problem-solving task (Klahr, 2000). It can also be viewed as a circular process, where research questions and hypotheses are formulated, investigations are planned and carried out, and evidence

is evaluated with regard to the hypotheses and the underlying theory (Mayer, 2007; Zimmermann, 2005). In order to achieve this circular process, various methodological and cognitive skills are inevitably important. Gott and Duggan (1998, p. 95), for example, mentioned the following skills that are needed to do science: ‘generate own ideas, hypotheses and theoretical models or utilize those postulated by others; design and conduct experiments, trials, test, simulations and operations; and evaluate the resulting data’. These skills can further be divided into cognitive skills, i.e. generate hypotheses, and methodological skills, i.e. conduct an experiment. These definitions also resonate well with the (AAAS, 1993) which asserts, “Scientific inquiry is more complex than popular conceptions would have it. It is, for instance, a more subtle and demanding process than the naive idea of ‘making a great many careful observations and then organizing them.’ It is far more flexible than the rigid sequence of steps commonly depicted in textbooks as ‘the scientific method.’ It is much more than just ‘doing experiments,’ and it is not confined to laboratories. More imagination and inventiveness are involved in scientific inquiry than many people realize, yet eventually strict logic and empirical evidence must have their day. Individual investigators working alone sometimes make great discoveries, but the steady advancement of science depends on the enterprise as a whole” (p. 9).

Elsewhere in the world, organizations and research committees like the OECD are advocating for the notions of critical thinking skills, problem solving skills, and creativity as major components of the modern-day skills that are required by students. Among these are the science process skills (SPS), also referred to as scientific method, scientific thinking and critical thinking. These skills are defined as a set of broadly transferable abilities, appropriate to many science disciplines and reflective of the behavior of scientists (Padilla, 1990). The science process skills are grouped into two types; basic and integrated. The basic (simpler) process skills provide a foundation for learning the integrated (more complex) skills. Basic SPS include observations, inferring, measuring, communication, classification and making predictions while integrated SPS consist of controlling of variables, formulating hypotheses, interpreting data, experimenting and formulating models (Padilla, 1990), which interchangeably same as scientific inquiry.

The terms scientific inquiry, scientific reasoning skills have been used interchangeably by some authors (Padilla, 1990; Wenning, 2007; Mayer 2007). In his study, Padilla (1990) found that experimenting ability, one of the integrated SPS, is closely related to the formal thinking abilities

described by Piaget. Thus, the instrument used to measure the SI is underpinned by Piaget's formal thinking operation that students at this level are supposed to have acquired.

2.5.1 Inquiry skills in science education

The inquiry is a term used both within education and in daily life to refer to seeking explanations or information by asking questions. It is sometimes equated with research, investigation or search for truth. Within education, Harlen (2013) posits that inquiry can be applied in several subject areas such as history, geography, arts, science and mathematics and so forth. When questions are raised evidence is gathered and possible explanations are considered. In each area, different kinds of knowledge and understanding emerge. This poses the question then, what distinguishes scientific inquiry from general inquiry?

Although both general and scientific inquiry leads to knowledge and understanding of the phenomena, scientific inquiry leads us to understand and explain the natural and synthetic world through direct interaction with the world and through the generation and collection of data for use as evidence in supporting explanations of phenomena and events (Harlen, 2013). Therefore, teachers are encouraged to adopt and make use of the approaches emphasized in the inquiry-based science education (IBSE) if they are to instill a culture of exploring, experimentation, investigations, and explanations within the schoolchildren. If students' inquiry skills are enhanced properly then their scientific reasoning skills are enhanced as well, the same is true for the scientific reasoning as well.

An inquiry-based science education means students are progressively developing key scientific ideas through learning how to investigate and build their knowledge and understanding of the world around them. Students use skills employed by scientists such as raising questions, collecting data, reasoning and reviewing evidence in the light of what is already known, drawing conclusions and discussing results. This learning process is all supported by an inquiry-based pedagogy, where pedagogy is taken to mean not only the act of teaching but also its underpinning justifications (Harlen, 2013). Inquiry-based learning is not an easy option, but OECD (2013) asserts that it is worth implementing because it promotes the understanding and development of skills needed by students to meet the demands of 21st century conditions. Current researches on science education (Zimmerman, 2013; Csapó, 2012; Harlen, 2013) echoed these sentiments that science education should enable students to develop key science concepts (big ideas) which enable

them to understand the events and phenomena of relevance in their current and future lives. Therefore, an inquiry is by no means a new concept in education, it is based on the recognition of children's active roles in developing their ideas and understanding. The studies of Piaget and the arguments of Dewey among others in the first half of the 20th century drew attention to the important role in their learning of children's curiosity, imagination and urge to interact and inquire (Harlen, 2013).

2.5.2 How to develop the scientific inquiry skills

What mechanisms can be used to stimulate and enhance students' scientific inquiry skills and by extension all of their reasoning skills? The development of scientific inquiry, as with the development of any reasoning, must necessarily be a slow and organic process in which the students construct the reasoning for themselves (Adey & Csapo, 2012). Morris, Croker, Masnick, and Zimmerman (2015) concurred that effective scientific skills require both deductive and inductive skills. Individuals must understand how to assess what is currently known or believed, develop testable questions, test hypotheses, and draw appropriate conclusions by coordinating empirical evidence and theory.

Furthermore, lessons which promote scientific inquiry skills provide plenty of opportunities for social construction (Adey & Csapo, 2012), that is to say, students are encouraged to talk meaningfully to one another, to propose ideas, to justify them and to challenge others. Recent research (Harlen, 2013) has shown that the adoption and the use of inquiry-based science learning has the potential to inculcate the scientific reasoning and thinking skills required in the 21st century. Harlen (2013) further posits that embracing inquiry-based science education recognizes its potential to enable students to develop the understandings, competencies, attitudes, and interests needed to exist in societies that are increasingly dependent on the application of science.

Notwithstanding that inquiry leads to knowledge of the particular objects or phenomena investigated, but more importantly, it helps build broad concepts that have wide explanatory power, enabling new objects or events to be understood (Harlen, 2013). A stimulating classroom environment is characterized by high-quality dialogue, modeled and organized by the teacher. This would require that students work within the zone of proximal development (ZPD) as proposed by Vygotsky (1978). The more knowledgeable students will be able to help their peers without the

peer feeling less important (Vygotsky, 1978). However, despite the over-whelming evidence that asking higher-level, open-ended questions has the potential to promote students' higher level reasoning and problem-solving, teachers still struggle to use these types of questions when interacting with their students (Gillies, Nichols, Burg, & Haynes, 2014). Therefore, the development of general scientific abilities is crucial to enable science students to successfully handle open-ended real-world tasks in their future careers (Bao et al., 2009). Bao et al., (2009) further state that teaching goals in science education include fostering content knowledge and developing general scientific abilities.

To implement these aspects of scientific inquiry in school, governments worldwide have set standards or benchmarks for science education. These documents have some common features concerning scientific inquiry (National Research Council, 1996). These standards form a conceptual framework for teaching science. They include more detailed standards and objectives for each subject (for biology/life science, and physical science {chemistry and physics}) that specify which aspects should be taught in these respective classes (Bybee, 2009). In Namibia, the aims and objectives of the National syllabus for sciences state, "learning experiences in the natural scientific area aim at increasing the learners' knowledge and understanding of the physical and biological world of which they are a part. This includes understanding how people use the natural environment to satisfy human needs, and how the environment may be changed in ecologically sustainable ways. Critical thinking, investigating phenomena, interpreting data, and applying knowledge to practical (experimental and investigative) skills and abilities are essential to understanding the value and limitations of natural scientific knowledge and methods, and their application to daily life" (Ministry of Education (MoE)., 2010, p. 2). The application of scientific knowledge and attitudes to health is of special relevance for the individual, the family, and society as a whole. These set standards foregrounded scientific inquiry as one area of competence for the three science subjects. For each of these subjects, however, there is a description of scientific inquiry and some examples of which aspects should be taught.

The methods used for scientific investigations play an important part in scientific inquiry. Mayer (2007) mentions the following methods: observing, investigating, describing, comparing, classifying, experimenting and using models. These descriptions are similar to the objectives in the national syllabi for science subjects in Namibia. It also resonates well with what is in the national broad curriculum of education in Namibia and the country's *Vision 2030*, which sees

Namibia as “developing from a literate society to a knowledge-based society, a society where knowledge is constantly being acquired and renewed, and used for innovation to improve the quality of life. A knowledge-based society requires people who are healthy, well educated, skilled, pro-active, and with a broad range of abilities” (MoE, 2010). Zimmerman (2005) describes activities for scientific investigations such as designing experiments, using apparatus and observing. Models can be used as an instrument for scientific investigations. Gilbert, Pietrocola, Zylbersztajn, and Franco (2000) claim that models and modeling should have a major impact on the learning of science in school.

2.6 Inductive reasoning

The third skill to assessed in this study is inductive reasoning (IR). It is one of the most widely researched areas in the literature because of its strong relation to fluid intelligence, problem solving and scientific reasoning skills. This is another area of interest for me with regard to the level Namibian students attain in this domain. It has been defined as a general thinking skill related to almost all higher order cognitive skills and processes (Csapó, 1997). There is no universally accepted definition of IR, though several definitions have been proposed (Molnár, Greiff, & Csapó, 2013). A classical understanding of IR is that IR is the process of moving from the specific to the general (Sandberg & McCullough, 2010). That is to say, IR is described as the generalization of single observations and experiences in order to reach overall conclusions. The IR test used in this study is based on Klauer’s (1990) definition of IR as discovering regularities by detecting similarities, dissimilarities, or a combination of both, with respect to attributes or relations to or between objects (Csapó, Molnár, & Nagy, 2014). The basic definition of inductive reasoning as explained in the (Multimedia Grolier's Encyclopedia, 1994, p. 287) asserts that,

“induction is a major kind of reasoning process in which a conclusion is drawn from particular cases. It is usually contrasted with deduction, the reasoning process in which the conclusion logically follows from the premises, and in which the conclusion has to be true if the premises are true. In inductive reasoning, on the contrary, there is no logical movement from premises to conclusion. The premises constitute good reasons for accepting the conclusion”.

I concur with the above explanations that the premises in inductive reasoning are usually based on facts or observations. There is always a possibility, though, that the premises may be true

while the conclusion is false, since there is not necessarily a logical relationship between premises and conclusion. Inductive reasoning is used when generating hypotheses, formulating theories and discovering relationships, and is essential for scientific discovery.

Induction can further be defined as the process whereby regularities or order are detected and, inversely, whereby apparent regularities, seeming generalizations, are disproved or falsified (Csapó, 1997). This is achieved by finding out, for instance, that all swans observed so far are white or, on the contrary, that at least one single swan has another color. To put it more generally, one can state that the process of induction takes place by detecting commonalities through a process of comparing. However, with inductive reasoning it is not enough to compare whole objects globally to each other. Instead, they have to be compared with respect to their attributes or to the relations held in common. That is the reason why all inductive reasoning processes are processes of abstract reasoning.

Literature have asserted that inductive reasoning is a basic component of thinking, and it is one of the most broadly studied procedures of cognition (Csapó, 1997). The teaching methods of instructions such as teaching by giving examples, questions and answers are considered to be the earliest methods of instructions, that enhances inductive reasoning. In addition, induction, or rather its role in generating scientific knowledge is one of the most enduring problems of philosophy (Csapó, 1997). Although recent cognitive research has resulted in a vast body of new information about inductive reasoning and has changed the understanding of its fundamental role, school instruction is far from using it effectively and several issues can be attributed to this less attention paid to the functions of induction in learning.

Additionally, inductive reasoning is usually contrasted to deductive reasoning: “Induction means establishing, deduction means applying rules” (Shye, 1988, p. 308). Thus, inductive reasoning enables one to detect regularities, rules, or generalizations and, conversely, to detect irregularities. We structure our world in this one way. It seems useful at the outset, to distinguish between inductive reasoning and inductive inferring. Inductive reasoning is aimed at detecting generalizations or regularities. If, for instance, a number of objects is given and if it is found that all of them are toys made of wood, a generalization or regularity has been discovered (Klauer et al., 2002). Should we extend this generalization to the totality of toys by stating that all toys are made of wood, then we would have made an inductive inference, although a false inference in this case. An inductive inference extends the generalization beyond the scope of experience by

asserting something about a non-observable universe of objects. Inductive reasoning, however, is confined to the observation at hand. It discovers regularity and order within a given set of objects.

There is consensus among researchers that inductive reasoning forms a central aspect of intellectual functioning. Ever since Spearman's (1923) study there has been no doubt about the close relationship between inductive reasoning and intelligence. Inductive reasoning is usually assessed by tasks consisting of classifications, analogies, series, and matrices (Goldman & Pellegrino, 1984; Sternberg & Gardner, 1983; van de Vijver, 1991). Figure 2, provides an analogy of the definition of inductive reasoning. Many intelligence tests contain one or more subtests of these varieties so that the contribution of inductive reasoning to intelligence test performance is beyond doubt.

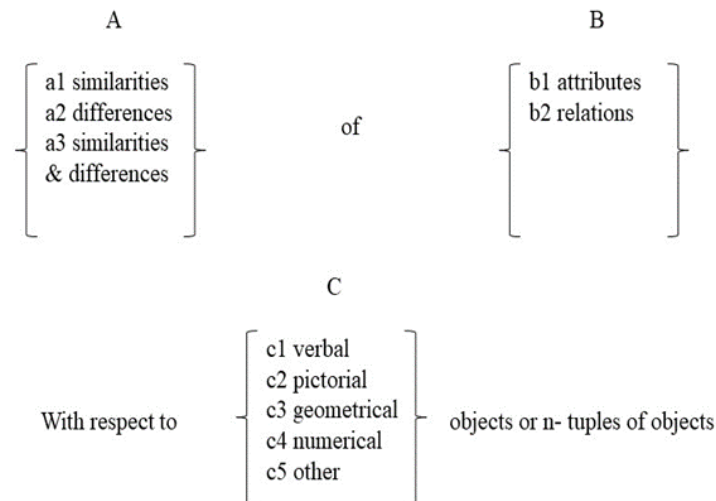


Figure 2 Definition of Inductive reasoning (Klauer, 1996, p. 38).

Thus, researches have suggested that, inductive reasoning relates well with domain specifics, such scientific inquiry and school achievement. Furthermore, inductive reasoning should be enhanced and taught explicitly in schools so as to stimulates the students' intelligence (de Konig, 2000).

2.7 Motivation to learn science

From an educational psychology perspective, in order to foster reasoning skills in science context effectively, students' reasoning and thinking processes need to be measured and monitored regularly alongside the exploration of students' motivational background of learning sciences (Glynn et al., 2011). The social cognitive theory perspective highlights that, the motivation of students to learn science in schools and colleges courses needs to be examined (Glynn et al., 2011) in order to determine the extend students are motivated to learn science. Social cognitive theory, developed by Bandura (1986, 2001, 2006) and extended by others (e.g. Pajares & Schunk, 2001; Pintrich, 2003), construes human functioning as series of reciprocal interactions among personal characteristics, environmental contexts, and behaviours. In social cognitive theory, students' learning is viewed as most effective when it is self-regulated, which occurs when students understand, monitor, and control their motivation and behaviour, leading to desirable learning outcomes. Motivation is defined in this theory as an internal state that arouses, directs, and sustains goal-oriented behaviour. By extension, the motivation to learn science patterns this process. Motivated students achieve academically by engaging in behaviour such as question asking, advice seeking, studying, and participating in classes, labs, and study groups (Schunk, Pintrich, & Meece, 2008).

Druger (2006) argues that one of the most important goals of an instructor of introductory college science courses is to help students become motivated self-learners. Like many science instructors, he evokes a variety of motivation components when expressing this goal: "We want students to enjoy science, recognize its role in the world, gain greater self-confidence about learning science, and want to learn more about science" (p. 39).

It is noteworthy that no single component captures the essence of what instructors, such as Druger, mean when they describe students who are motivated to learn science. That is because the motivation to learn, as conceptualized in social cognitive theory, is a multi-facets construct. The motivation components and attributes were reviewed by researchers such as (Eccles & Wigfield, 2002; Glynn & Koballa, 2006; Koballa & Glynn, 2007; Pintrich, 2003; Schunk et al., 2008). Examples of these components are *intrinsic motivation*, which deals with the inherent satisfaction in learning science for its own sake (e.g. Eccles, Simpkins, & Davis-Kean, 2006); *self-determination*, which is about the control students believe that they have over their learning of science (e.g. Black & Deci, 2000); *self-efficacy*, which refers to students' belief that they can

achieve well in science (e.g. Lawson, Banks, & Logvin, 2007); and *extrinsic motivation*, which involves learning science as a means to a tangible end, such as a career or a better grade (e.g. Mazlo et al., 2002). These mutually supporting components of motivation contribute positively to the arousal, direction, and sustainment of students' science-learning behaviour. Together, these components constitute a componential model of motivation derived from social cognitive theory (Glynn et al., 2011).

When measuring the motivation to learn science, science education researchers attempt to determine why students strive to learn science, what emotions they feel as they strive, how intensively they strive, and for how long they strive (Bandura & Locke, 2003). Measuring the motivation to learn science is challenging because a construct and its components are not directly observable variables. For this reason, they are called latent variables. Although latent variables cannot be directly observed, they can be measured by means of observed variables (items) that serve as empirical indicators. The items on the Science Motivation Questionnaire were designed to serve as empirical indicators of components of students' motivation to learn science in elementary school science through university courses (Glynn et al., 2011).

Motivation to learn science benefits young students who aspire to be future scientists. But, just as importantly, motivation to learn science benefits all students by fostering their *scientific literacy*, which is the capability to understand scientific knowledge, identify important scientific questions, draw evidence-based conclusions, and make decisions about how human activity affects the natural world (OECD, 2007). The importance of all students becoming scientifically literate is advocated internationally (Feinstein, 2011; Kelly, 2011; Roberts, 2007).

2.8 Technology-based assessments

Current researchers such as Adey and Csapó (2012) suggested ways of assessing reasoning skills and inquiry skills. They argued that computerized testing can be closer to the ideal individual interview than a paper and pencil assessment advocated by earlier research, such as Piaget's clinical interviews. Furthermore, administering the same test to every subject improves the objectivity of the assessment (Adey & Csapó, 2012). The advantages of technology-based assessment cannot be overemphasized. The administration of tests to large sample at a one go, automatic scoring and immediate feedbacks are just but some of the usefulness of technology-based assessment.

The new skills of the 21st century are not derived from the educational standards of the 20th century (Mayrath, Clarke-Midura, Robinson, & Schraw, 2012). As we live in a new world after the evolution of computer technologies. Educational practices need to shift from the industrial revolution to the knowledge revolution, so are the assessments also need to be different (Mayrath et al., 2012). Meanwhile, it is a great source of concern whether tests, teachers, and policies are moving quickly enough to keep pace. The above statement is summarized briefly by (Mayrath et al., 2012, p. 40) that:

“When I went to school, I did not receive any training or experiences with 21st century skills. There were no rooms for multiparty games that required timely opportunistic communication and negotiation strategies with invisible players. Collaborative problem solving to achieve group goals was not part of our curriculum. I never learned how to manage limited resources and understand tradeoffs between factors with an interactive simulation. We primarily lived in a world of print in books rather than a rich colorful world of visualizations and multimedia. I never was encouraged or taught how to ask deep questions (why, how, what if, so what) and to explore novel hypotheses because all of our curriculum and subject matters was preplanned by the teacher”.

This statement confirms what most people in the new information technology environment experience because platforms such as Google, blogs, and quick electronic access to millions of information sources that vary from hidden to free association to rigorously validated wisdom did not exist then.

Interestingly, some of the current trends and aspects such as inquiry skills and reasoning skills needed in the 21st century are also conspicuously absent from most of current Namibian curricula, standards, and high stakes assessments in grade 10 and 12 respectively in the Namibian context. “Our society continues to concentrate on assessing 20th century reading, writing, mathematics, science, and academic knowledge, typically with multiple choice and other psychometrically validated tests that can be efficiently administered in one or a few hours” (Mayrath et al., 2012, p. 8). The landscape of skills and knowledge being tested does not stretch to the new environment. The testing format does not sufficiently tap the functional cognitive procedures and social practices of today. Many groups are trying desperately to correct this misalignment, both inside and outside of the assessment industries, but the process is slow and laborious, with politics complicating everything (Csapó, Ainley, Bennet, Latour, & Law, 2012). Furthermore, technologies are currently available to perform fine-grained formative assessments

of knowledge and skills over long stretches of time that can be measured in months or years. If Namibian education system could just up its game, the technology could do a great deal in education fraternity. Technology can be used to track everything the student reads from digital media in both formal and informal learning environments over the course of a year, which could not be possible if teachers have to do that manually. Imagine administering the standard psychometric practice of crafting multiple choice questions to assess reading comprehension ability in a one-hour, anxiety-ridden, high stakes test or a two hours scientific reasoning test using paper and pencil methods. There needs to be a more rapid but rigorous methodology for assessing students in the 21st century. Thus, easy to use assessment instruments are necessary for everyday school practice. Testing students using paper-based tools require a large number of human resources and the immediate feedback is hard to realize. Technology-based assessment may provide feasible solutions due to its innovative characteristics such as innovative item design, automated coding, feedback and data processing (Csapó et al., 2012).

2.9 The need for Technology-Based Assessment (TBA)

We live in an ever-changing world – demographic change, rise of automation and workforce structural change, globalization, and corporate change are some major driving forces that demand fundamental transformations in education and skills on an individual level (Bao & Koenig, 2012). There has been a rise in studies that emphasize the importance of technology-based assessment. As Namibia is a relative newcomer in the international arena, technology-based assessment is almost nonexistent. However, a government policy on information communication and technology (ICT) exists on paper, but the implementation thereof is lagging behind. Modern mechanisms for measuring the performance of the education standards in Namibia against international benchmarks are yet to be established below grade 12 end of school year examination (Ministry of Education, 2007). This means that the assessments of scientific reasoning, inquiry skills and critical thinking online have not yet been established.

To effectively measure and assess current trends and issues in the 21st century, nations are encouraged to adopt the use of technology in classroom and as assessment media. Teaching and learning program should include skills such as SR, SI and IR as they are economic imperative, and assessment is a fundamental component of any pedagogical program (Mayrath et al., 2012). Advances in assessment theory, educational psychology, and technology create an opportunity to

innovate new methods of measuring students' skills with validity, reliability, and scalability (Mayrath et al., 2012). "We are not using the full flexibility and power of technology to design, develop, and validate new assessment materials and processes for both formative and summative uses" (U.S. National Educational Technology Plan, March, 2010). Educational organizations around the world are also calling for 21st Century Skills (21CS) and looking to technology as a means to improve learning, motivation, and collaboration. Research supports the use of technology for content delivery (Mayer, 2001; Mayrath, Nihalani, & Robinson, 2011; Csapó et al., 2012). However, only over the last ten to twelve years has a research base been established that systematically explores the use of technology-based assessments to measure complex Knowledge, Skills, and Abilities (KSAs) (Bennett, 1999; Csapó, Molnár & Nagy, 2014; Molnár, Greiff, & Csapó, 2013). The above statement is true for most of the developed world and at least they have begun somewhere, and this should serve as a guide and motivation for the third world countries such as Namibia to follow suits. Modern research on cognition and human learning, combined with emerging technologies, offers new possibilities for teaching and assessing higher-order thinking skills. However, systematic research is needed to determine how best to measure these complex KSAs using technology-based assessments. To sum up, it is argued that, technology has had a significant effect on educational assessments in a number of dimensions (Csapó et al., 2012). For example, it is found to improve the efficiency of collecting and processing of data, it enables the sophisticated analysis of the available data, supports decision-making and provides rapid feedback to participants and stakeholders alike. However, in as much TBA has proven to be an efficient way of operating there are still challenges facing some countries, especially in the developing world. These challenges include the provision of internet infrastructures, computers, tablets and all technologically based equipment to schools among others.

2.10 Socio economic status (SES)

As for students' socio-economic status, a number of different indicators have been used to categorize SES. International studies (e.g. PISA) use a number of different background variables. As the impact of these may differ from country to country, cross-country comparisons are more valid if a broad array of variables is used (Kuger, Klieme, Jude, & Kaplan, 2016). Out of these variables, in PISA a complex index of economic, social and cultural status (ESCS) is composed (OECD, 2016). As usually strong correlations are found between the background variables, for

national assessments, fewer variables suffice to meaningfully represent students' SES. For example, in Hungary, the type of settlement did not have a unique contribution to the variance in students' achievements (Nikolov & Csapó, 2018). Although there are large differences between learners' achievements in villages and cities, these are fully explained by the differences in the educational level of the students' parents living in these two environments. Therefore, the single best SES variable explaining most of the variance in students' achievements has been their mothers' level of education (Csapó, 2010). In developing countries like Namibia, we do not have available data about the SES and the students' achievement.

2.11 Summary

This chapter reviewed literature SR, SI and IR. Several researchers highlighted the importance of and the roles played by these skills (SR, SI & IR) in the students' education in the 21st century. Literature revealed that the methods used for scientific investigations play an important part in the scientific reasoning and scientific inquiry. The following methods of observing, investigating, describing, comparing, classifying, experimenting and using models were highlighted as the core of SR and SI. Similar descriptions can also be found in the Namibian National Educational Broad Curriculum Standards for Science (MoE, 2010). Gilbert et al. (2000) claim that models and modeling should have a major impact on the learning of science in school.

Recognizing the importance of inductive reasoning has led a number of international researchers to investigate its relationship with other constructs and background variables in educational contexts (Csapó, 1990, 1997, Csapó, Lorincz, & Molnár, 2012; Molnar, Greiff, & Csapó, 2013). Their results confirmed the significance of inductive reasoning in human cognition. Measurements on different samples showed the rapid development of the skill in the early years of schooling, thus, it is vital that interventions should be implemented in this period. Findings also underpin the phenomenon that different forms of inductive reasoning are strongly connected and the most significant component was analogical reasoning. Analogical reasoning helps to understand something that may be true more like inductive reasoning rather trying to establish a fact as in deductive reasoning. Analogical reasoning is useful while learning to understand new information or even while engaging in persuasive arguments. As such, this kind of reasoning is highly beneficial in various behaviors such as decision making, problem solving, enhancing memory, explanation, perception, communication, emotion and creativity. Strong correlations

were found between tests measuring application of knowledge in science, mathematical understanding and foreign language. Literature also showed that no significance differences in performances between genders but achievements are strongly influenced by parents' level of education.

In terms of TBA, the benefits of technology-based assessments cannot be overemphasized. Csapó et al. (2012) outlined the long-term advantages of Information Communication Technology (ICT). The literature further reveals that ICT enables traditional assessments processes to be carried out more efficiently. It provides frequent and precise feedback for participants in learning and teaching that cannot be achieved by any other means.

CHAPTER 3 AIMS OF THE RESEARCH AND STRUCTURE OF THE EMPIRICAL STUDIES

3.1 Research aims and structure of the empirical studies

The purpose of this study is to assess students' abilities in thinking and reasoning skills such as, scientific reasoning, scientific inquiry and inductive reasoning skills in Namibia. We also wished to explore the possibilities of an online assessment in Namibia. This study connects three rapidly developing areas of educational research and places them in the context of the development of the Namibian education system. First, improving the conditions of science education in order to attract more young people to the Science, Technology, Engineering, and Mathematics (STEM) professions is a goal of many countries as the supply of young professionals graduating in these areas does not meet the demand of modern economies. One of the main directions in this area is the quality of science education, especially its contribution to the improvement of students' higher order thinking skills. Second, educational assessment is receiving growing attention both in research and in practice. If certain psychological constructs are made measurable, it opens a path for conducting precise training experiments, while feedback provided by the assessment may orient practice. Third, testing is transferred to a technological basis making even more constructs measurable, while reducing the costs and timeframe of the assessments.

This study consists of four empirical sub studies (1) online assessment of scientific reasoning, inductive reasoning and motivation to learn science, the aims of this study were to explore the possibilities of online assessment, to investigate students' knowledge and reasoning skills based on log file analysis, and to explore the relationship between reasoning skills in a science context and the motivation to learn science in Oshana region, Namibia. (2) The possibilities of assessing students' scientific inquiry skills abilities using an online instrument, the aim of study 2, was to pilot the on-line instrument for the assessment of scientific inquiry skills of the 9th and 11th grade students. Primarily, the study wishes to construct a scale tapping different components of scientific inquiry to measure it broadly, economically and efficiently using an on-line assessment platform, Electronic Diagnostic Assessment (eDia). (3) Assessment of scientific reasoning, scientific inquiry and inductive reasoning with students' socioeconomic status. The relationship among the three main constructs (SR, SI, & IR) for this study was also explored, (4) and the fourth supplementary study was conducted to further explore if English reading comprehension skills affect students' achievement in SR and IR.

Generally, these four studies aim to assess students' abilities in thinking and reasoning skills such as, scientific reasoning, scientific inquiry and inductive reasoning skills in Namibia. We also wish to explore the possibilities of an online assessment in Namibia. Data collection procedures, timeline and steps are summarized in (Table 2).

Table 2. The timeline and the research activities

Timeline	Research activities	Instruments	Samples
September 2015	<ul style="list-style-type: none"> Contextualization of the research focus, writing and presenting research plan. Development and translation of assessment items. Literature search and reviews. 		
June-July 2016	<ul style="list-style-type: none"> Piloting the assessment instruments in the 	<ul style="list-style-type: none"> Scientific reasoning-(Korom et al., 2012; 2017) Inductive reasoning -(Pásztor et al., 2017) Science Motivation questionnaire-(Glynn et al., 2011) eDia platform 	<ul style="list-style-type: none"> 5th & 7th grades, N=616
September 2016	<ul style="list-style-type: none"> Data analysis and presentation of the results. 		
January 2017	<ul style="list-style-type: none"> Assessing the students' through an online platform. 	<ul style="list-style-type: none"> Scientific inquiry skills test-(Nagy et al., 2015; Korom et al., 2017) eDia platform 	<ul style="list-style-type: none"> 9th and 11th grades N=118
February 2017	<ul style="list-style-type: none"> Data analysis of the collected data in January 2017 and presentation of results to conferences and seminars. Organizing final assessment tools for the large-scale assessments. 		
June-August 2017	<ul style="list-style-type: none"> Large scale assessment of Scientific reasoning, Scientific inquiry and Inductive reasoning skills in the. 	<ul style="list-style-type: none"> Lawson Classroom Test of Scientific reasoning skills-(Lawson, 2000) Scientific inquiry skills- (as above) Inductive reasoning- (as above) Paper and pencil methods. 	<ul style="list-style-type: none"> 10th and 12th grades N=582
September 2017	<ul style="list-style-type: none"> Data analysis continues and presentation of results in conferences and seminars. 		
January/February 2018	<ul style="list-style-type: none"> Supplementary assessment of the students Reading comprehension skills, Scientific reasoning and Inductive reasoning. 	<ul style="list-style-type: none"> Reading comprehension skills tests - (Csapó & Nikolov, 2009; 2018) Scientific reasoning-(Lawson, 2000) Inductive reasoning- (as above) Paper and pencil methods used 	<ul style="list-style-type: none"> 8th grade students N=250
March 2018	<ul style="list-style-type: none"> Data analysis and results presentations. 		

3.2 Research questions and hypotheses

Research questions and hypotheses are divided according to the four sub studies that guided the research. More specifically, corresponding research questions and hypotheses for each sub study are as follow.

Research question for sub study 1

1. What are the psychometric properties of the SR and IR tests? Will the tests be reliable?
2. How well do grade five and seven students perform on scientific and inductive reasoning tests?
3. How does girl students' performance differ from boy students' in scientific and inductive reasoning?
4. How do the background variables (e.g. parents' level of education) influence their performance?
5. What are the relationships among the subtests of scientific reasoning skills?
6. What is the relationship between inductive and scientific reasoning, scientific reasoning and the motivation to learn science of the students?

Hypotheses for sub study 1

- H₁. The psychometric properties of the tests are acceptable.
- H₂. We expect the 7th graders to perform significantly better than the 5th graders.
- H₃. Based on the literature, we expect no significant differences between genders.
- H₄. In accordance with the literature, we hypothesized that students from high socio-economic backgrounds to significantly perform better than students from low economic backgrounds.
- H₅. We hypothesized medium correlations between the subconstructs.
- H₆. We expect strong correlation between SR and IR, and between SR and motivation to learn science (SMQII).

Research question for sub study 2

1. What are the psychometric properties of the scientific inquiry skills test?
2. How well do grade 9 and 11 students perform on the scientific inquiry skills?

3. Is there a significant difference in performance between genders on scientific inquiry test?
4. How do parents' levels of education influence students' performance?
5. What are the relationships among the subtests of the scientific inquiry skills test?

Hypotheses for sub study 2

- H₇. We expect the psychometric properties to be acceptable.
- H₈. We hypothesized significant achievement differences between the two grades
- H₉. We expect no significant differences in achievement between genders.
- H₁₀. Literature have emphasized that mothers' level of education influence students' achievement, therefore we expect high socio-economic status students to significantly perform better than low socio-economic status.
- H₁₁. We expect strong correlations between the subconstructs and the whole scale.

Research question for sub study 3

1. Are the psychometric properties of the tests acceptable in this sample?
2. How well do the 10th and 12th grade students perform on scientific reasoning, inductive reasoning and scientific inquiry skills? (What are the differences among the three tests within the grade?)
3. How do girls differ from boys on scientific reasoning, inductive reasoning and scientific inquiry skills?
4. How do the background variables (e.g. parents' level of education) influence their performance?
5. What is the relationship among the subtests of scientific reasoning, inductive reasoning and scientific inquiry skills?

Hypotheses for sub study 3

- H₁₂. We expect the psychometric properties to be acceptable.
- H₁₃. Significant differences in achievement between the two grades is anticipated.
- H₁₄. As per the literature, no significant differences between gender is hypothesized among the three constructs.
- H₁₅. Based on our previous research results, we hypothesized significant differences in achievements between low and high socio-economic students.

H₁₆. Based on our previous results, strong correlations are expected between sub constructs and whole scales

Research question for sub study 4

1. How well do the 8th graders perform in reading comprehension (RC), IR and SR?
2. What are the relationships between students' RC skills and their performance in IR and SR?
3. How does students' SES affect their performance in IR and SR?

Hypotheses for sub study 4

H₁₇. Based on our previous research results, we hypothesized better achievements in the tests.

H₁₈. We expect strong correlations between the three tests.

H₁₉. As with the previous studies, we expect significant differences in achievements between low and high economic status students.

CHAPTER 4 METHODOLOGY

4.1 Introduction

This study used a cross-sectional study design to assess the thinking and reasoning skills of the Namibian students in three domains, i.e. scientific reasoning, scientific inquiry skills and inductive reasoning skills. As mentioned in Chapter one, the assessment of thinking and reasoning skills are not a feature in Namibia, therefore the cross-sectional study was deemed appropriate for the following reasons; 1) “ a single snapshot of the cross-sectional study provides researchers with data for either a retrospective or a prospective enquiry, 2) a cross-sectional study can also bear several hallmarks of longitudinal studies of parallel groups (e.g. age groups) which are drawn simultaneously from the population” (Cohen, Manion, & Morrison, 2010, p. 201) .

4.2 Samples

The type of problems studied here do require the composition of a nationally representative sample; however, due to a limited time frame and the resources at my disposal, the ideal of composing a national sample could not be realized. Moreover, we needed samples that were large enough to bear the major typical characteristics of schooling in Namibia. Thus, samples were drawn from schools in Oshana, Omusati, Oshana and Oshikoto regions in the northern part of Namibia where 50 % of the population lives. Whole school classes were chosen for group testing, and we tried to achieve the best representation of schools in the area in terms of quality and type of schooling. Figure 3, shows the Namibian maps with its 14 political regions. The samples were drawn from the Omusati, Oshana, Oshikoto regions (2, 3, & 5) as denoted in the map, (see Figure 3).

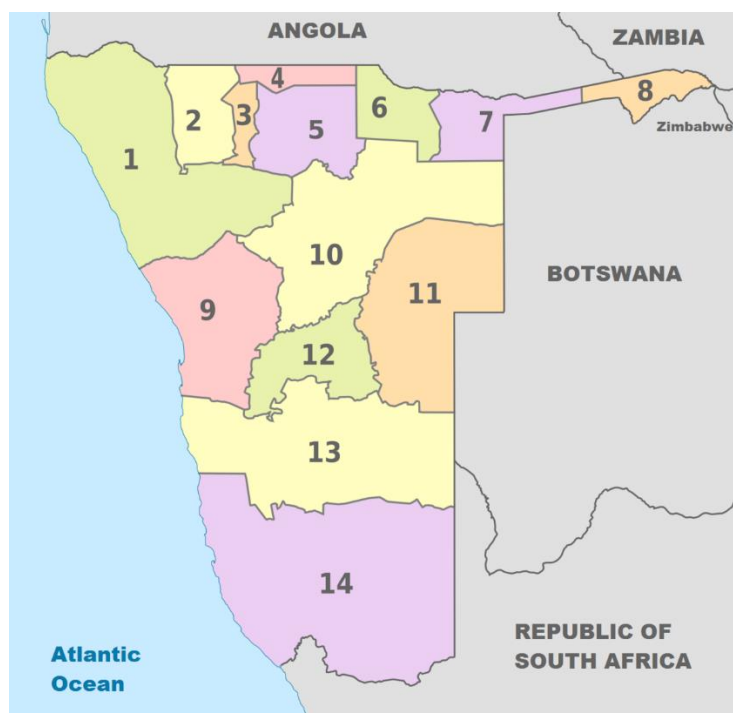


Figure 3 Namibian Regions Source: www.namibiansafari.com/natravelMAP1.htm

Primary schooling consists of a period of 8 years (schooling begins at the age of 6 years) and the schools use the same centrally developed curriculum and the same set of textbooks and are fairly homogeneous. Some primary school students (5th and 7th graders) took part in the first study (Sub study 1), and in order to achieve a representative sampling, samples were taken from the densely populated regions of Namibia and mostly from secondary schools (8th to 12th grades) were students come from different villages across the region to schools in the regional capital. At the time of the data collection, the secondary school curriculum in Namibia was the same, i.e. secondary schools start from 8th grade to the 12th grade, no other types of secondary schools such as grammar school or vocational schools existed. However, the education system was also going through a reform phase, and it is likely that in the near future there might be different type of secondary schools in Namibia (i.e. grammar schools, vocational schools). The three subsequent studies took place at the secondary school level, 8th, 9th & 11th and 10th & 12th graders respectively. The youngest age when our tests can be used is the 5th grade of primary school; below this age, reading difficulties were anticipated (Csapó, 1997). The oldest age group that can be tested within the educational system is the 11th grade and beyond. Between these two points, measurements took place. Further data on the samples composition are summarized in (Table 3).

Table 3. The samples of the study

Studies	Samples	Instruments
Study 1	Grades=5 & 7 N=616	SR- (Korom et al., 2012; 2017) IR-(Pásztor et al., 2017) SMQ- (Glynn et al., 2011)
Study 2	Grades =9 &11 N=118	Scientific inquiry skills-(Nagy et al., 2015; Korom et al., 2017).
Study 3	Grades=10 & 12 N=582	LCTSR-(Lawson, 2000) SI- (Nagy et al., 2015; Korom et al., 2017) IR-(Pásztor et al., 2017)
Study 4	Grade=8 N=250	RC-(Csapó & Nikolov, 2009; Nikolov & Csapó, 2018) LCTSR-(Lawson, 2000) IR-(Pásztor et al., 2017)

We also wished to find out if the background variable such as socio-economic status of the students influences their achievement in the three domains tested. One of the socio-economic indicators is the parents' level of education. The level of parental education is considered to be the best indicator for students' socio-economic background (Keller, Neumann, & Fischer, 2017). Table 4, 5, 6 and 7 show the categories of parents' level of education used in each study. The categories have been changed due to the results of the sub study one, for example we had school mature exam and college categories in the first study and realized that these categories were not clear to students and changed them in the subsequent studies.

Table 4. The distribution of the level of education of the students' parents for study 1.

Parents' level of education	Mothers %	Fathers %
1. Didn't finish elementary school	10.2	7.0
2. Elementary school	5.2	8.0
3. Vocational school	3.0	3.7
4. School mature exam	5.8	3.6
5. College	7.0	11.2
6. University	37.5	35.1
7. I don't know	30.0	27.4
8. Missing data	2.0	4.0
Total	100	100

Table 5. The distribution of the level of education of the students' parents for study 2

Parents' level of education	Mothers %	Fathers %
1. Did not go to school	2.5	3.4
2. Did not finish primary school	3.4	1.7
3. Primary education	8.5	5.9
4. Secondary education	47.5	30.5
5. Higher education	38.1	46.6
6. PhD degree	-	2.5
7. I do not know	-	9.3
Total	100	100

Table 6. The distribution of the level of education of the students' parents for study 3

Parents' level of education	Mothers %	Fathers %
1. Did not go to school	12.9	15.5
2. Did not finish primary school	5.8	4.0
3. Primary education	14.6	16.0
4. Secondary education	38.3	34.5
5. Bachelor degree	18.7	16.0
6. Master's degree	7.2	8.9
7. PhD degree	1.5	1.9
8. I do not know	0.9	3.3
Total	100	100


Table 7. The distribution of the level of education of the students' parents for study 4

Parents' educational level	Mothers (%)	Fathers (%)
1. Did not go to school	3.6	7.2
2. Did not finish primary school	5.2	5.2
3. Grade 10	9.6	7.2
4. Grade 12	30.8	25.6
5. Bachelor degree	18.4	18.0
6. Master degree	28.0	29.6
7. PhD degree	4.4	7.2
Total	100	100

4.3 Instruments

The scientific reasoning skills test and inductive reasoning as well as motivation to learn science questionnaires were used for sub study 1, where students are required to use their cognitive skills (scientific reasoning and general thinking skills inductive reasoning skills) see appendix A and B, to answer questions based on different sub-constructs of scientific processes. It assesses different thinking and reasoning skills essential for learning science and learning in general see sample tasks in (Figure 4, 5, and 6)

We pour the milk from the glass into the bowl. Which statement is true?
Click on the answer.




- ☐ Both the volume and the shape of the milk change.
- ☐ Only the volume changes, but not the shape of the milk.
- ☐ Only the shape of the milk changes, but not the volume.
- ☐ Neither the shape nor the volume changes.

☐ Previous ☐ Next

Figure 4 Sample task for conservation

Learners boiled 2 litres of water to 80°C in a pot. The temperature of water was 20°C at the beginning and it took 5 minutes to boil the water. How much time will they need to boil 2 litres of water to 80°C , if the temperature of water is 50°C at the beginning? Click on the answer.



☐ 2.5 minutes ☐ 5 minutes ☐ 7 minutes ☐ 10 minutes

☐ Previous ☐ Next

Figure 5 Sample task for proportional reasoning

Observe the figure. Decide whether the following statements are true or false.
Click on the answer.

```

graph TD
    Vascular[Vascular plants] --> NoFlowers[Plants without flowers]
    Vascular --> WithFlowers[Plants with flowers]
    NoFlowers --> Ferns[Ferns]
    WithFlowers --> Gymnosperms[Gymnosperms]
    WithFlowers --> Angiosperms[Angiosperms]
    Angiosperms --> Dicots[Dicots]
    Angiosperms --> Monocots[Monocots]
  
```

If a plant is a dicot, it will not have flowers. ☐ True ☐ False

Gymnosperms are vascular plants. ☐ True ☐ False

There are some flowering plants which are monocots. ☐ True ☐ False

All ferns are flowerless. ☐ True ☐ False

☐ Previous ☐ Next


Figure 6 Sample task for logical operations

As it can be seen from the tasks, figures 4, 5, and 6, the SR tests contain cognitive tasks in science context, which require students to use their cognitive knowledge to solve the tasks. Furthermore, most of the tasks also demand a certain level of reading comprehension in order to understand what is asked. In sub study 2, scientific inquiry skills test (see appendix C) was used, for both sub study 2 and 3. The tasks are domain specific and students are required to use the inquiry skills and knowledge to solve the tasks. The demands for science content is necessary here. Again, as in any other tests, a certain level of reading comprehension is required see sample tasks in Figure 7.

Ningini wanted to investigate whether the temperature of water affect the amount of sugar dissolved in it. She poured equal amount of water into two identical cups. She then heated one cup to 20° C and the other cup one to 60° C. She put one sugar cube into each cup, and stirred until the cube dissolved. She repeated this process until cubes did not dissolve anymore. She counted how many sugar cubes were dissolved in each cup.

Which of the following hypothesis did she test? Click on it.

- As the amount of sugar increases, so does the amount of dissolved sugar increase.
- As temperature of water increases, so does the amount of dissolved sugar increase.
- As the amount of water increases, so will the amount of dissolved sugar increases.
- Stirring will increase the amount of dissolved sugar.







Previous Next

We observe how light and water affect the growing of plants. We assume that plants need light and water too for staying alive. How should we choose the conditions/parameters to verify our statement? Make the experimental parameters.

Put the name of the environmental factor to the observations.

There is light There isn't light There is water There isn't any water

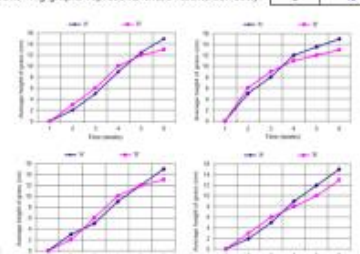
Experiment 1	Experiment 2	Experiment 3	Experiment 4
			

Previous Next

Maria and Hilya compared the rate of growth of two different types of grass. They planted same number of grass seeds at the same time in two identical pots with same amount of soil. They kept the pots under same conditions for six weeks and made observations at the same time each week. They recorded average height of grass in each pot. They recorded their observations in the table.

Which of the following graphs represents these results correctly? Click on it.

Time (weeks)	Average height of grass (cm)	
	A	B
1	0	0
2	2	3
3	5	6
4	9	9
5	12	10
6	15	13



Previous Next

Students compared the different states of matter of water. The ice tray was filled with water to the top (brim). The tray was put in a freezer. The ice cubes formed were higher than the top (brim) of the tray. They made the following table from their result.

Experiment	Activity	Result
1	They filled completely the ice tray, then put it in the freezer.	Ice cubes overflowing out from the ice tray.
2	Ice cubes are put into water.	Ice cubes floated on the surface of water and their size was continuously decreasing.

Could we make the following conclusions from these experiments? Click on the right answer for each conclusion.

The freezing point of water and the melting point of ice are the same. Yes No

If water is frozen, its volume will change. Yes No

Ice cubes melt at room temperature. Yes No

Ice has smaller density than water. Yes No

Previous Next

Figure 7 Sample tasks of the SI tests

Informed by the results and findings of sub study 1 and 2, all the three domains were then administered to students from the 10th and 12th grades respectively. In all the sub studies, background variables such students' socio-economic status was included. The scientific reasoning test used in sub study 1, was replaced in sub study 3, and 4, due to its low reliability, by Lawson Classroom Test of Scientific Reasoning (LCTSR), 2000 version, see sample tasks in Figure 8, and more detailed one in appendix D. As it can be seen in the sample tasks, figure 4-6 and Figure 8, the SR and the LCTSR, tasks are asking almost the same cognitive dimension. As already stated, students need to apply their reasoning skills to solve the tasks problems. The SI and IR test remained the same throughout the studies.

1. Suppose you are given two clay balls of equal size and shape. The two clay balls also weigh the same. One ball is flattened into a pancake-shaped piece. Which of these statements is correct?

- a. The pancake-shaped piece weighs more than the ball
- b. The two pieces still weigh the same
- c. The ball weighs more than the pancake-shaped piece

2. *because*

- a. the flattened piece covers a larger area.
- b. the ball pushes down more on one spot.
- c. when something is flattened it loses weight.
- d. clay has not been added or taken away.
- e. when something is flattened it gains weight.

3. To the right are drawings of two cylinders filled to the same level with water. The cylinders are identical in size and shape.

Also shown at the right are two marbles, one glass and one steel. The marbles are the same size but the steel one is much heavier than the glass one.

When the glass marble is put into Cylinder 1 it sinks to the bottom and the water level rises to the 6th mark. *If we put the steel marble into Cylinder 2, the water will rise*

- a. to the same level as it did in Cylinder 1
- b. to a higher level than it did in Cylinder 1
- c. to a lower level than it did in Cylinder 1

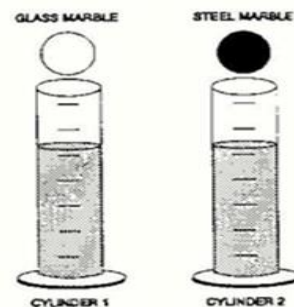


Figure 8 Sample tasks of the LCTSR tests

Furthermore, the Hungarian based scientific reasoning; scientific inquiry skills and inductive reasoning tests were adapted and used in the Namibian context. The Magyar Tudományos Akademia (MTA-SZTE) - Research Group on the Development of Competencies of the Institute of Education, University of Szeged, develops these tests online. The test assesses Hungarian students' reasoning and general thinking skills (Adey & Csapó, 2012; Csapó, 1997; Korom et al., 2017; Korom et al, 2012; Pásztor, Molnár, Korom, B. Németh, & Csapó, 2017). For the scientific reasoning (see appendix A), items were developed based on the Lawson classroom test of scientific reasoning skills framework model (Lawson, 2000). The Lawson model was simplified to accommodate the basic skills required in the school science curricula. Korom et al. (2017), designed a simplified model that consists of five sub-constructs, which are conservation of volume and matter, proportional reasoning, correlational reasoning, probabilistic reasoning, and classification. The items are intended to measure the reasoning skills of primary school students through the secondary school science curriculum and beyond.

The scientific inquiry tests (see appendix C and sample items Figure 7) items were developed based on Wenning's (2007) scientific inquiry skills framework model. The Wenning model was simplified to accommodate the basic skills required in the school science curricula. (Nagy, Korom, Pásztor, Veres, & Nemeth, 2015), designed a simplified model that consists of seven sub constructs. The sub constructs are: setting research questions, hypothesis identification, identification of variables, variable planning, experimental plans, data handling technique and making conclusions. Again, the items are intended to measure the inquiry skills of the primary to secondary school science curriculum. Tasks require students to apply their reasoning skills and recall the experiments/practical work/investigations projects they have done at school from primary through to the 12th grade.

Although several tests for inductive reasoning (see sample tasks Figure 9) were described in the literature, a new IR test was developed by the (MTA-SZTE research group) (Pásztor, 2016; Pásztor et al, 2017).

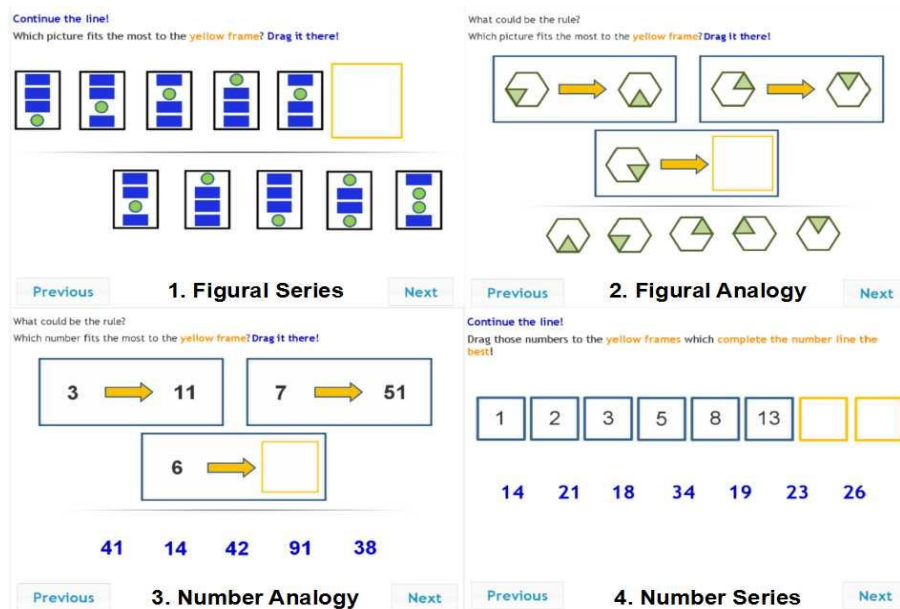


Figure 9 Sample items (Inductive reasoning) used in all sub studies.

This test was adapted and used in Namibian context as well. Research has shown that IR test are context free and can be applied almost in any cultural setting (Csapó, 1997). (see appendix B and sample items in figure 9). Literature further suggested that (1) verbal tests may translate poorly; and (2) the tests were prepared for further use in other surveys and training experiments where sensitive and reliable measurement instruments are required, therefore I adapted the IR test of four

subconstructs from the MTA-SZTE research group (Pásztor et al., 2017). The test on inductive reasoning skills measure different thinking skills essential for learning in general and learning science in particular. The IR test is based on Klauer's (1990) definition of IR as discovering regularities by detecting similarities, dissimilarities, or a combination of both, with respect to attributes or relations to or between objects (Klauer, 1990; Klauer et al., 2002). The inductive reasoning test administered in Namibia consisted of 56 items. The test consisted of four subtests, which were figural series and figural analogy (Pásztor & Molnar, 2015), number analogy, number series (Csapo, 1997). The texts of the SR used in sub study 1, SI and IR were translated from the Hungarian language into English. Furthermore, a supplementary study was also carried out to gauge the students' ability in reading comprehension and see if it influences their achievement in SR and IR.

Since the test items (SR, IR, and SI) were developed in a Hungary and written in Hungarian language, a professional translator, with the help of people that were involved in the task development translated the texts from Hungarian to English. Tasks that were deemed unsuitable in the Namibian context were removed and some were adapted. Before implementation, the tasks were also sent to two colleagues from the Department of Mathematics, Science, and Sport Education at the University of Namibia for suggestions. Two experienced science teachers from Namibia also reviewed the translated texts. Table 8 summarized the instrument used in this study.

Table 8. Instruments used in the study

Studies	Grades	Instruments	Methods	Descriptions
Study 1	5 & 7	SR- (Korom et al., 2012; 2017) IR-(Pásztor et al., 2017) SMQ- (Glynn et al., 2011)	eDia. platform	SR- tasks used here are closer to the interpretation of scientific reasoning as different thinking processes in science context. IR-domain general and cultural free content reasoning tasks. SMQ-questionnaires about motivation to learn science in schools
Study 2	9 & 11	Scientific inquiry skills- (Nagy et al., 2015; Korom et al., 2017).	eDia. platform	SI-tasks based on the different stages of scientific investigations. Tasks are science content embedded.
Study 3	10 & 12	LCTSR-(Lawson, 2000) SI- (Nagy et al., 2015; Korom et al., 2017) IR-(Pásztor et al., 2017)	Paper and pencil	LCTSR- tasks are more complex, and general than the SR. It contains tasks that are closer to scientific inquiry as well. SI & IR used here are the same as described above in study 1 & 2.
Study 4	8	RC-(Csapó & Nikolov, 2009) LCTSR-(Lawson, 2000) IR-(Pásztor et al., 2017)	Paper and pencil	RC-tasks ranged from single words expression, sentence to short passages. Focus was on meaning and not form. LCTSR & IR used here were the same as above.

4.4 Procedures

4.4.1 Data Collection

Informed by the current trends in research, the online data collection was carried out through the Electronic Diagnostic Assessments (eDia) platform (Csapó & Molnár, 2017) for study 1 and study 2. Students were ferried from their schools to the University of Namibia's ICT rooms, due to non-functional ICT equipment and poor internet connections at their schools. Immediate feedback was given after task completion. The administration of the tests took approximately 90 minutes. For study 1, scientific reasoning and inductive reasoning tests were administered to the 5th and the 7th grade students during the Namibian winter (June/July) 2016. When the results showed that the tests were a bit difficult for the 5th and 7th graders, a decision was made to move to a much older sample, i.e. to secondary school phase. For study 2, the scientific inquiry skills test was administered to the 9th and 11th grades samples during the Namibian summer of January/February 2017. Although the data were collected at different times, the measurements took place within two years, and we do not know of any relevant differences between the two testing periods. Therefore, we analyzed the data as if they were the results of simultaneous cross-sectional measurements. The same procedures of group testing were applied in each case.

The results of the online testing proved to be acceptable, however due to technical problems, mainly the lack of internet and ICT facilities at school, the data collection for Study 3 & 4 were carried out using the paper and pencil methods. The online results of scientific reasoning (study 1) informed our decision to adapt and use the validated and popularly used Lawson Classroom Test of Scientific reasoning (LCTSR), (Lawson, 1978, 2000) for study 3 and 4. The reliability indices of the whole test were within the acceptable value (Cronbach alpha= .74) but at the subscales level the indices were really low. This proves that the Hungarian version of SR, needs major revision and improvement before it could become a validated research instrument. However, the other two tests, i.e. scientific inquiry and inductive reasoning (Pásztor, Molnár, Korom, Németh, & Csapó, 2017), proved to be good, hence the continuation with them to study 3. The data collection took place in the winter (June/July) of 2017. We decided to keep to the secondary school phase and focused on the two grades that write the external examination in Namibia, 10th and 12th grades. Out of curiosity, and the results of study 1, 2 and 3, further informed us, language issues were also explored. So, study 4 was envisaged to explore whether English reading comprehension skills would have an effect on the students' tests achievement. The paper and pencil method was used again and an online version of the reading test (Nikolov & Csapó, 2018) was converted into paper and pencil format. This data collection was done in January/February 2018 (summer in Namibia).

4.4.2 Data analysis

As stated in Chapter one, the main goal of this study was to assess Namibian students' abilities in thinking and reasoning skills in the three domains of scientific reasoning, scientific inquiry and inductive reasoning skills. The aim was also to explore the possibilities of an online assessment, but the results of study 1 & 2 showed that it would not work. The data were analyzed accordingly using the Statistical Package for Social Sciences (SPSS), Mplus and IRT.

Independent sample t-tests were used to find the differences in performance between the grades and between genders. Analysis of variance (ANOVA) was applied to examine the differences in students' achievement according to their parents' level of education (socio economic status). Furthermore, item response models (IRT) was used as it is in line with the research goal of determining the students' ability in three domains of thinking and reasoning. "The main idea of item response theory (IRT) is to use a mathematical model for predicting the probability of success

of a person on an item, depending on the person's 'ability' and the item 'difficulty' (Adams & Wu, 2002, p. 28). Typically, the probability of success on an item for people with varying ability is plotted as an "item characteristic curve" (ICC). Item response models typically apply a mathematical function to model the probability of a student's response to an item, as a function of the student's "ability" level (Adams & Wu, 2002).

Structural equation modeling (SEM) (Bollen, 1989) was also used to test the underlying measurement model for scientific reasoning, scientific inquiry and inductive reasoning skills. All measurement models were computed with Mplus. Maximum Likelihood Squares and Mean- and Variance-adjusted (MLSMV) estimation was used (Muthén & Muthén, 2010). Different fit indices, such as the Tucker–Lewis Index (TLI), the comparative fit index (CFI) and the root mean square error of approximation (RMSEA), were computed to assist in determining model fit. Nested model comparisons were conducted using a special chi square (χ^2)-difference test for the MLSMV estimator (Muthén & Muthén, 2010). Apart from that, a simple bivariate correlation analysis was applied to find the relationships between and among subtests and between the main construct.

4.5 Summary

In this chapter, the processes involved in gathering the data that would answer the research questions are described. In any research study it is important to use the appropriate design and methodology to achieve credible and reliable research outcomes (Cohen et al., 2007, p. 183). On the basis of the questions posed for the study, it was considered appropriate to use a quantitative, cross-sectional design. Likewise, quantitative data gathering tools such as test taking and questionnaires were found to be the most relevant to assess and ascertain students' ability level in scientific reasoning, scientific inquiry and inductive reasoning. The process of choosing test instruments and how the instruments were adapted is explained. A synopsis of how the data were treated and analyzed is briefly juxtaposed as the detail data analysis processes are explained under each sub studies. In the next chapter, the results and discussions that emerged from the four sub studies are outlined.

CHAPTER 5 THE STUDIES

5.1 Sub study 1. Online assessment of scientific reasoning, inductive reasoning skills and motivation to learn science among the Grade 5 & 7 students in Oshana Region

5.1.1 Introduction

In this study, the possibilities of online assessment of scientific reasoning and the motivation to learn science were explored in the Oshana region, Namibia. The Government of the Republic of Namibia recognizes education as one of the key inputs for economic development, human welfare, collective progress, and environmental protection. To this end, the ICT policy was introduced in 2001. “The purpose of this policy is to prepare all Namibia’s learners, students, teachers, and communities of today for the world economy of tomorrow” (Ministry of Basic Education, Sport, and Culture [MBESC], 2001). This policy has also been long envisaged, as the statement appearing in the country’s ‘Vision 2030’ which states that; “As we move towards a knowledge-based development paradigm, as stipulated in Namibia’s Vision 2030 ‘integrating ICT education and training into education and training systems, issues of access to the local and global pool of knowledge and information become paramount” (MBESC, 2001, p. 6).

In order to develop scientific reasoning and thinking skills effectively, we need to explore how some factors such as students’ motivation to learn science and the socio-economic status of the students influence their performance. To date many studies have been conducted on assessing students’ scientific knowledge mostly in developed countries, for example, the large-scale international assessments, PISA and TIMSS (Bao et al., 2009; Bybee & Fuchs, 2006; Mayer et al., 2014). Namibia does not take part in international assessment programs and large-scale studies are rarely done in Namibia.

5.1.2 Methodology

5.1.2.1 Participants

The sample of the study was drawn from the fifth and seventh graders (N=621; 268 boys; 348 girls; age M=12.40, SD=1.19) from five different schools in Oshakati and Ongwediva towns. For grade 5, the sample was N=275 (121 boys; 152 girls, age M=11.19, SD=.68) and the grade 7 were 346 (147 boys; 196 girls, age M=13.23, SD=.61). The schools were selected based on the

availability of ICT infrastructure at the school, therefore the sample is not representative; typically, students with above-average social backgrounds attend these schools. The entire 5th and 7th grade students that were present during the day of the test took part in the project. Due to the selection of the schools, the sample consisted of a number of students whose parents have a high level of education (see table 9). The level of parental education is considered as the best indicator for students' socio-economic background (Keller, Neumann, & Fischer, 2017). However, ANOVA showed no significant difference between parents' level of education and students' achievement.

Table 9. The distribution of the level of education of the students' parents

Educational level	Mothers (%)	Fathers (%)
1. Didn't finish elementary school	10.2	7.0
2. Elementary school	5.2	8.0
3. Vocational school	3.0	3.7
4. School mature exam	5.8	3.6
5. College	7.0	11.2
6. University	37.5	35.1
7. I don't know	30.0	27.4
Missing data	2.0	4.0
Total	100	100

5.1.2.2 Instruments

As described in Chapter 4, this study was based on scientific reasoning skills, an inductive reasoning test where students were required to use their cognitive skills (scientific reasoning) to answer questions based on different sub-constructs of scientific processes. It assessed different thinking and reasoning skills essential for learning science and learning in general.

The online assessment tool consisted of 16 tasks (36 items) assessing different thinking skills in science context (Korom et al., 2017), such as conservation, proportional reasoning, correlational reasoning, probabilistic and classification skills as well as working with logical operations, see sample tasks on Figure 4, 5, and 6, (Korom et al., 2012, 2017). The test based on Klauer & Phye's (2008) inductive reasoning was also used (see figure 9). The Science Motivation Questionnaire II (SMQ, Glynn et al., 2011) was also used, an online version was developed of this tool, (see table 10).

5.1.2.3 Procedures

The online data collection was carried out through the eDia platform (Csapó & Molnár, 2017) in the University of Namibia's ICT department. Immediate feedback was given after task

completion. Students were ferried from their schools to the University of Namibia's ICT department as their ICT equipment was not working and the internet connections at their schools was poor. The administration of the test took approximately 60 minutes. In terms of SMQII, students answered a questionnaire online with the following instructions: 'In order to better understand what you think and how you feel about your school science subjects, please respond to each of the following statements from the perspective of (When I am in a science class...) see table 10. Response are on 5-point Likert scale: (Never; Rarely; Sometimes; Usually; Always). We managed to collect questionnaire data from 165 students.

Table 10. The SMQII questionnaire that was developed into an online version (Glynn et al., 2011, p. 1165)

Components (Scales and Items)	Never (1)	Rarely (2)	Sometimes (3)	Usually (4)	Always (5)
Intrinsic Motivation					
The science I learn is relevant to my life					
Learning science is interesting					
Learning science makes my life more meaningful					
I am curious about discoveries in science					
I enjoy learning science					
Self-efficacy					
I am confident I will do well on science tests					
I am confident I will do well on science labs and projects					
I believe I can master science knowledge and skills					
I believe I can earn a grade of "A" in science					
I am sure I can understand science					
Self-determination					
I put enough effort into learning science					
I use strategies to learn science well					
I spend a lot of time learning science					
I prepare well for science tests and labs					
I study hard to learn science					
Grade Motivation					
I like to do better than other students on science tests					
Getting a good science grade is important to me					
It is important that I get an "A" in science					
I think about the grade I will get in science					
Scoring high on science tests and labs matters to me					
Career Motivation					
Learning science will help me get a good job					
Knowing science will give me a career advantage					
Understanding science will benefit me in my career					
My career will involve science					
I will use science problem-solving skills in my career					

This material was used with the full permission as indicated here "Science educators who wish to use the Science Motivation Questionnaire II, have the permission to do so" (Glynn et al., 2011, p. 1165).

5.1.3 Results

5.1.3.1 Psychometric properties

The reliability index of the scientific reasoning skills test was acceptable (Cronbach-alpha=.70) for the whole sample, however at the subscales level it yielded such low-reliability indices that it could not be interpreted. Table 11 shows the reliability indices per each grade, with the 5th grade yielding a lower Cronbach alpha=.64 compared to the 7th grade Cronbach alpha=.69 (see table 11). For the science motivation questionnaire, the reliability (Cronbach alpha= .91) was quite high for the whole sample and per grade. More details on the reliability indices are shown in table 11. Furthermore, the scientific reasoning skills test was moderately hard for the students: $M=37.83\%$; $SD=13.34\%$ (see table 12). Students performed quite well in the proportional and correlational subtests compared to the rest of the subtests (see table 12).

Nonetheless, positive correlations were found between the subtests and the scientific reasoning achievement, also indicated in Table 12. Strong positive correlations are observed between conservation of mass and volume and proportional reasoning subscales with the main scale (scientific reasoning). However, the correlational reasoning scale showed a weak correlation with the main scale.

5.1.3.2 Grade differences

The performance between the two grades (Table 11) was statistically significant ($t(616)=7.87, p<.01$). This means that scientific reasoning skills developed with age, as the 7th graders performed significantly better than the 5th grade students. However, in respect of reliability, the test behaved the same in both grades, as there is no large difference between the Cronbach's alphas. These results could suggest the need to improve the reasoning and thinking skills of the students at the primary school level.

Table 11. The reliability indices and descriptive statistics of the test and the questionnaire

Tests	No of items	Cronbach's alpha			Mean% (SD%)		
		Grade 5	Grade 7	Both grades	Grade 5	Grade 7	Both grades
Scientific reasoning skills (Korom et al., 2017)	36	.64	.70	.69	34.3 (12,3)	40.6 (13.5)	37.8 (13.3)
Science motivation questionnaire (Glynn et al., 2011)	25	.90	.91	.91	2.84 (.76)	3.06 (.71)	2.92 (.75)

Furthermore, the findings indicate that the test might be too hard for the primary school students. Table 12, shows more detailed descriptive statistics for the SR in both grades.

Table 12. Mean scores and correlations between subtests and main construct (scientific reasoning).

<i>Scales</i>	Both grades (N=621)				Grade 5 (N=275)				Grade 7 (N=346)			
	<i>M</i> (%)	<i>SD</i> (%)	<i>r</i>	<i>p</i>	<i>M</i> (%)	<i>SD</i> (%)	<i>r</i>	<i>p</i>	<i>M</i> (%)	<i>SD</i> (%)	<i>r</i>	<i>p</i>
Scientific reasoning	37.83	13.34			34.39	12.34			40.56	13.46		
Conservation of mass & volume	35.07	20.22	.80	.001	32.23	18.51	.76	.001	37.32	21.23	.83	.001
Proportional	41.01	19.81	.63	.001	37.13	19.64	.63	.001	44.09	19.43	.61	.001
Correlational	43.00	35.52	.34	.001	38.00	35.76	.32	.001	46.97	34.86	.33	.001
Probabilistic	36.67	28.05	.45	.001	33.36	27.49	.48	.001	39.31	28.26	.42	.001
Classification	38.44	18.61	.50	.001	34.75	17.55	.49	.001	41.37	18.94	.46	.005

One-parameter Rasch analyses were also carried out in order to gain a more detailed picture about the test. The EAP/PV reliability was about .70, which is acceptable. Further investigation showed that few items were suitable for differentiating students at low skill levels as shown in (figure 10). The analysis revealed that there were no items in a low ability level especially in the 5th grade, and several items at the top with no student capable of getting a correct score. This means that the test was a too difficult for the students as the same trend is observed in the 7th grade. However, the distribution in 7th grade was a little better than the 5th grade, as reported in the first section (grade differences) where 7th graders performed significantly better than the 5th grade. Few students were at the lower end of the distribution, and few were positioned at the high ability items on the top.

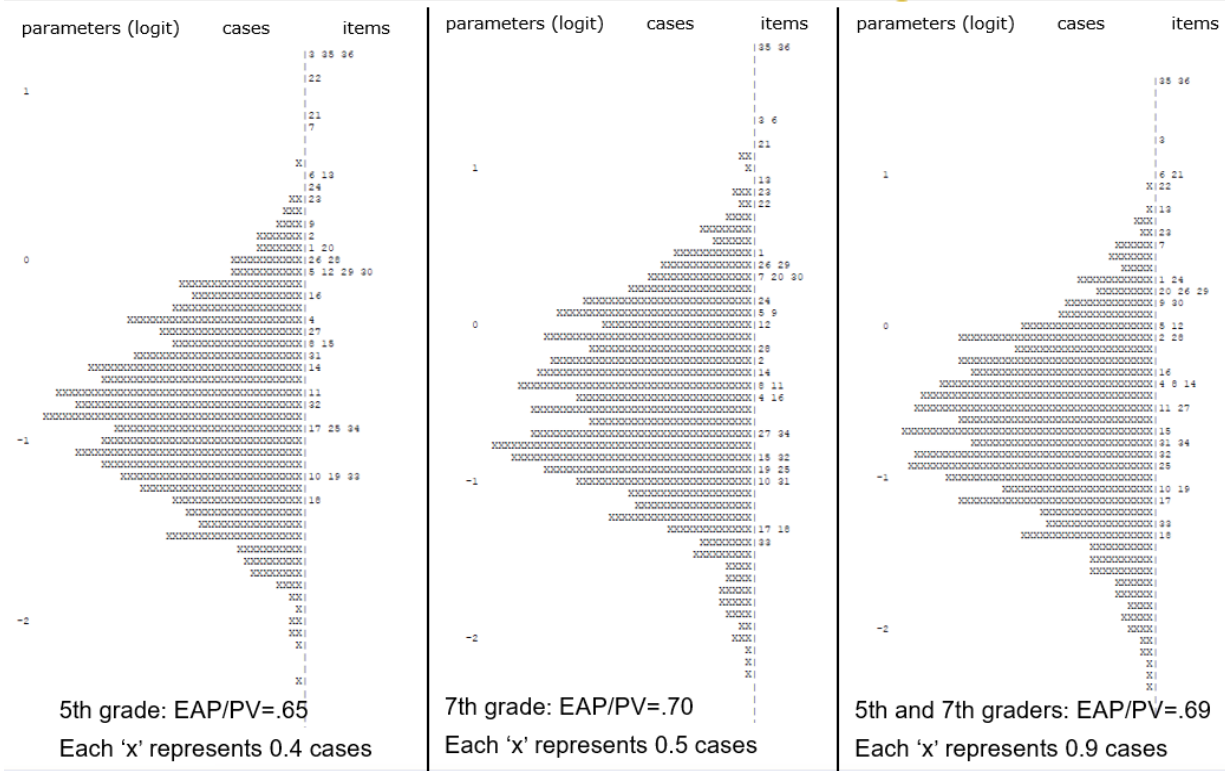


Figure 10 Person-item maps for the reasoning skills test

The reliability of the science motivation questionnaire (SMQ) was good at the subscales level and at the main scales level (Cronbach-alpha=.91 main scale) given the fact that this is an already validated instrument available in the literature (table 13). Average scores were relatively high ($M=3.06$, $SD=.71$), thus, students reported that they are motivated to learn science. However, the lowest score was found in the areas of intrinsic motivation and self-determination, (see table 13).

Table 13. Descriptive statistics and reliability indices of SMQ

Scales	No. of items	Both grades (435)			Grade 5 (N=270)			Grade 7 (N=165)		
		α	M	SD	α	M	SD	α	M	SD
SMQII	25	.91	2.92	.75	.90	2.84	.76	.91	3.06	.71
Intrinsic motivation	5	.65	2.78	.85	.64	2.72	.85	.65	2.85	.86
Self-efficacy	5	.78	3.04	.89	.76	2.94	.90	.82	3.19	.85
Self determination	5	.72	2.84	.86	.71	2.76	.89	.73	2.98	.79
Grade motivation	5	.72	3.06	.84	.70	2.97	.85	.75	3.19	.82
Career motivation	5	.74	2.96	.88	.70	2.87	.89	.80	3.10	.86

Note: The scores ranged from 0 to 4.

Table 14, shows the correlations between SMQ and the scientific reasoning results. Except for self-efficacy with no significant correlations ($p<.05$). The weak correlations between scientific

reasoning and motivation with the relatively high average scores on motivation suggest that regardless of students' scientific reasoning skills' level they still intend to learn science in general. This may provide a promising basis for further instructions.

Table 14. SMQ correlations with scientific reasoning

	Both grades (N435)		Grade 5 (N=270)		Grade 7 (N=165)	
<i>Scales</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
SMQII	.23	.001	.20	.001	.21	.008
Intrinsic motivation	.20	.001	.16	.006	.19	.011
Self-efficacy	.19	.001	.18	.003	.14	.070
Self determination	.17	.001	.15	.017	.16	.041
Grade motivation	.22	.001	.21	.001	.20	.011
Career motivation	.22	.001	.22	.001	.17	.035

5.1.3.3 Log results

The task analysis on conservation of volume and matter see for example figure 4, revealed that 70.2% of the students could not answer the question correctly, 22.3% of students thought that both volume and the shape of the milk would change, 28.6% thought that only the volume would change and 12.1% thought that neither the shape nor volume would change. This result shows that the students have a serious problem with understanding the concept of volume. The same trend continues with the task on proportional reasoning see figure 5; about 78% of the participants could not give the correct answer. According to the log analysis, many who failed had the idea that to solve the task, one chose to multiply by two or divide by 0.5 but students mixed up the methods (40.5%). For instance, on Figure 6, only 19.9% of students were able to solve all four logical operation tasks correctly. It has to be noted that no explicit science knowledge is necessary to solve this task. One only needs to use thinking and reading comprehension skills to solve it. Thus, the low achievements also reveal fundamental obstacles in science learning as students seem to have major problems understanding basic science figures.

5.1.3.4 Inductive reasoning and scientific reasoning

The reliability of the domain general thinking skills was good compared to the scientific reasoning. The analyses showed a (Cronbach alpha = .80) and all four subscales showed acceptable internal consistency of (Cronbach alpha > .70), which is the accepted index for the ability tests (Muthén & Muthén, 2010). Further details are shown in table 15. Moreover, analyzing the samples

performance on both scientific and inductive reasoning revealed that 7th grade performance was significantly higher than the 5th grade, see table 15 and figure 11.

Table 15. Reliability indices and descriptive statistics of the inductive reasoning

Tests/subtests	Both grades (N=621)				Grade 5 (N=276)			Grade 7 (N=345)		
	No. of items	α	M (%)	SD (%)	α	M (%)	SD (%)	α	M (%)	SD (%)
<i>Inductive reasoning</i>	55	.80	31.73	12.10	.79	27.77	11.42	.78	35.00	11.72
Figure series	15	.75	40.11	22.29	.78	36.74	22.91	.73	42.80	21.43
Figure analogy	15	.74	36.19	21.10	.70	28.62	18.85	.72	42.24	20.86
Number analogy	14	.74	32.77	17.94	.74	31.57	17.48	.74	33.73	18.26
Number series	11	.71	12.91	17.15	.80	9.55	16.16	.72	16.00	17.45

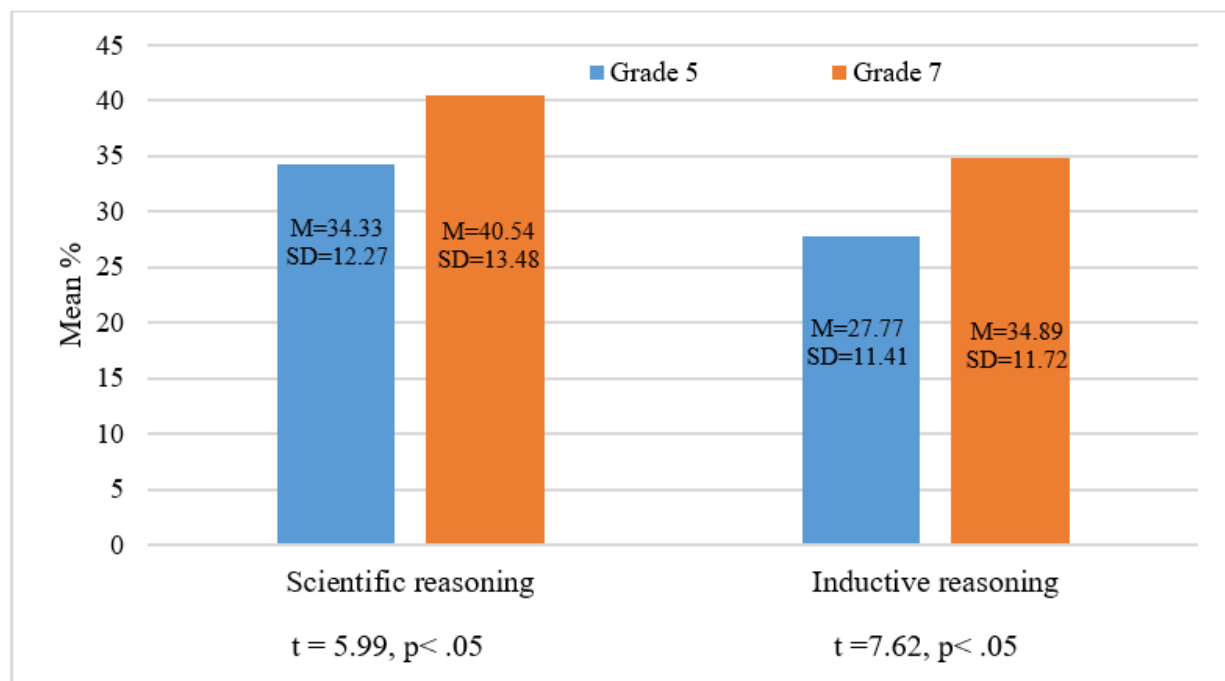


Figure 11 The 5th & 7th grades performance on scientific and inductive reasoning

Furthermore, the frequency analysis also showed that students did not do well in number analogy and number series subscales. This raises the issue of whether the Namibian education system prepares students well in basic numeracy and scientific education. The distribution of the achievements revealed that none of the students achieved above 80% and only three students scored below 10% (see Figure 12). Figure 12 also revealed that about 40 % of the 5th grade only managed to score between 20-29 %. Very few 5th grade students were able to score more than 60 %. However, the analysis further indicates that 20 % of the 7th grade students scored above 50 % in both scientific and inductive reasoning.

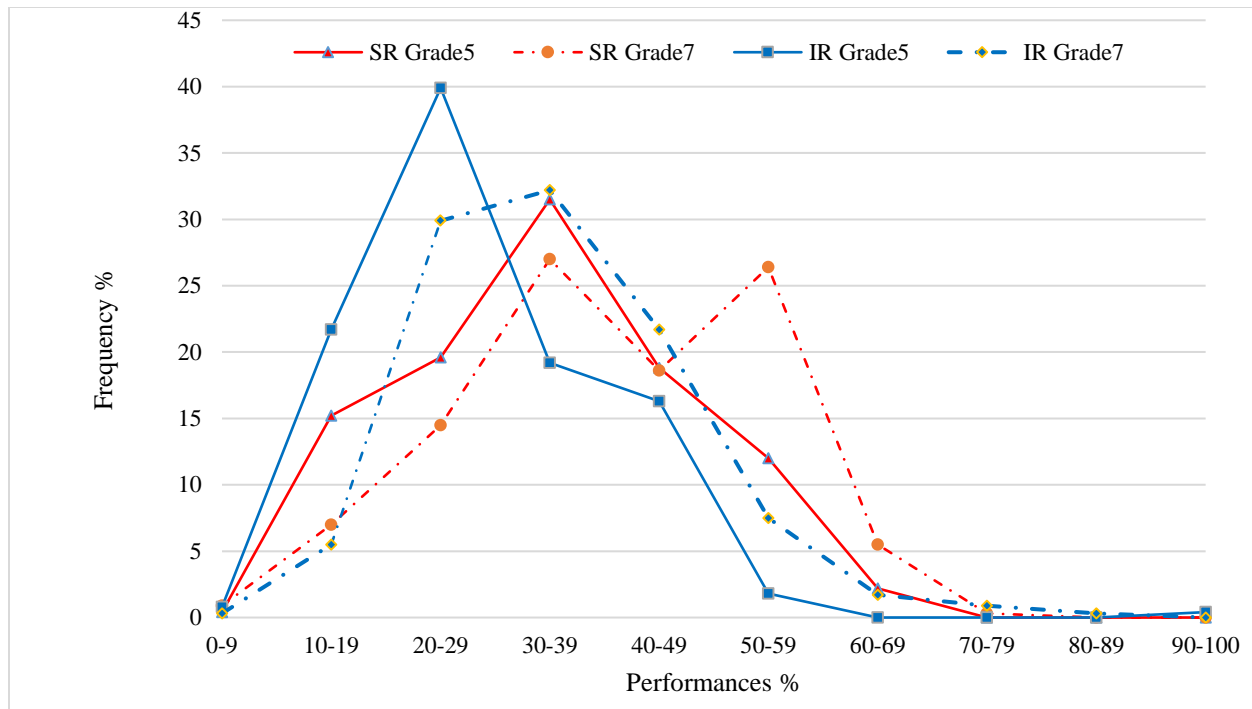


Figure 12 Students' performance on scientific reasoning and inductive reasoning

The correlation between scientific reasoning and domain general inductive reasoning was stronger in the 7th grade, see table 16. Within subtests, the strongest correlation was found with figural analogies in both grades. However, very low correlations were found between number analogy in grade 5 and number series subscales with scientific reasoning in grade 7 compared to the rest.

Table 16. Correlations between inductive reasoning and scientific reasoning

Variable	Grade 5	Grade 7	Both grades
Inductive reasoning	.38**	.40**	.44**
Figure series	.27**	.29**	.32**
Figure analogy	.34**	.40**	.42**
Number analogy	.14*	.24**	.23**
Number series	.25**	.22**	.27**

**Correlation is significant at the .01 level (2-tailed). *Correlation is significant at the .05 level (2 tailed)

Linear regression analysis in table 17, showed that inductive reasoning explained almost 17 % of variance ($F=19.45$, $p<.001$) in the 7th grade and just about more than 15 % of variance ($F=16.96$, $p<.001$) in the 5th grade. Only figural analogies have significant beta coefficient in both grades.

Table 17. Effect of subscales of inductive reasoning on scientific reasoning

Grade	Independent Variable	B	SE	Beta	t	p	r	Explained variances %
5	Figure series	.07	.05	.10	1.51	.13	.26	2.6
	Fig analogy	.16	.05	.24	3.55	.00	.34	8.16
	Number analogy	.09	.07	.08	1.31	.19	.14	1.12
	Number series	.17	.08	.13	2.12	.04	.25	3.25
7	Figure series	.06	.04	.09	1.43	.16	.29	2.34
	Figure analogy	.20	.04	.31	4.75	.00	.40	12.04
	Number analogy	.08	.06	.08	1.33	.19	.24	1.92
	Number series	.03	.06	.03	.46	.64	.22	.66

Dependent Variable: Scientific reasoning

5.1.4 Discussions

The online assessment instruments for SR proved to be moderately reliable despite the low mean achievement, therefore hypothesis 1 has been proven. This may mean that the assessment is best suited to an older age group. While the advantages of technology-based assessment, such as online test administration and automatic calculation of scoring, reduced the time and cost of the testing (Pásztor, Molnár, & Csapó, 2015), process cannot be overemphasized, the findings revealed that technology-based assessment via online testing may not be viable in the Oshana region, Namibia at the moment. This is because the researchers had to ferry students from their schools to the University of Namibia (UNAM), computer labs for the data collection to take place. The sample was taken from urban schools and none of the schools had a functional computer laboratory, one can just imagine the situations in villages where most of the population lives. However, care was taken not to generalize these findings as they may not be representative.

Apart from literature, indicating that online assessments provide teachers with an easy-to-use instrument for monitoring the development of students' thinking and may contribute to the development of effective teaching methods, at the moment the suitable assessment instrument for Namibia would be paper and pencil. According to figure 8, the SR showed that more items proved difficult for the students at this level as there were quite a number of items that could not be answered. Comparing some items qualitatively with the primary science syllabus, the enhancement and promotion of reasoning skills are hardly emphasized. As it stands, one may conclude that the curriculum focuses more on content and is exam driven. The effort to include 21st-century skills

appears to be minimal. The Namibian learners need to acquire these 21st-century skills; the question is, is our education system ready and equipped to impart them? However, the analysis revealed that the scientific reasoning test must be revised in order to carry out more reliable assessments in this age cohort in Namibia. New items have to be added with some items needing to be modified and deleted in order to find a fit between the test and the students' skills specially to increase the discrimination power of the test in low skill levels. However, the current test versions may be piloted in older age groups. No significant differences in gender performance was found just as in line with the literature. Furthermore, as expected, grade 7 students perform significantly better than the grade 5 students. Examining the relationship between thinking skills' achievement and some aspects of motivation to learn science, we found weak correlations between scientific reasoning and motivation with the relatively high average scores on motivation which may suggest that regardless of students' scientific reasoning skills' level their general intent is to learn science. This may provide a promising basis for further instructions.

This sub study 1, was one of the first attempts to carry out an online assessment in Namibia at primary school level. The results indicate that technology-based assessment may provide schools and teachers with a user-friendly instrument for monitoring the development of students' reasoning skills. However, my experience suggests that many public schools in Namibia do not have a functional ICT infrastructure in order for the online assessments to be an everyday school practice.

In addition, the psychometric analyses of the reasoning skill tests revealed that the tests were challenging for this age cohort in Namibia and should be revised in order to achieve more reliable assessments in this age group. The validity and reliability of the SR test could not be fully established as it yielded low reliability indices both at the whole scale and sub scales level. However, the IR reliability both at the whole scale and sub scales level were high. Research has also confirmed that IR is not affected by testing mode or cultural issues as long as culture free tasks are used. Therefore, the IR test validity was confirmed, except for the poor performance of the students, which suggests that these students may need training in IR.

Descriptive statistics and IRT analyses of the SR, revealed that the test was problematic for the primary school students; the suggestion is that this instrument (SR) needs further development to achieve more reliable assessment. Lawson SR test could be used instead, to assess the reasoning and thinking skills at the secondary school level. The task analysis revealed that

students were not able to solve simple tasks of conservation, proportional and logical operations. This performance of the tasks highlights the need to train and expose students explicitly to different types of reasoning and thinking skills. This finding also suggests that teachers' training should be adapted in order to understand how to incorporate these skills into their everyday teaching.

It should also be mentioned that the sample was biased towards students with good social backgrounds and one wonders what would be the case if the same tasks were administered in the primary schools in and around the rural areas where one finds most students are from low socioeconomic background. Concerning the non-significant correlations found between self-efficacy and test performance, this might be attributed to language. Self-efficacy might be a new word to most primary school students. The weak correlations between thinking skills in science and motivation and the relatively high average scores on motivation suggest that regardless of students' thinking skills' level students still intend to learn science. This may provide a promising basis for further investigation.

So far, the validity and reliability issues have been established. The results of sub study 1 proved that the SR test under construction needs further improvement for it to be a valid instrument to use. It is also worth noting here that this test even in Hungary was and is actually under construction before it is validated. Due to the limited time I had available, the decision was then made to use the already validated instrument of SR (Lawson, 1998, 2000). The IR test proved to be reliable, hence the decision was made to retain it and use it for future assessments. In the next section, (sub study 2), the newly designed test of scientific inquiry was employed (Korom et al., 2017).

5.2 Sub study 2. The possibilities of assessing students' scientific inquiry skills abilities using an online instrument: a small-scale study in the Omusati Region, Namibia

5.2.1 Introduction

The aim of this study was to assess students' scientific inquiry skills using an online test instrument in the Omusati region, Namibia. The demands on learners and thus, education systems are evolving fast (Csapó & Funke, 2017). In the past, education was about imparting knowledge, now it is about making sure that students develop a sustainable ingenuity and the navigation skills to find their own way through an increasingly uncertain, changing world (Csapó & Funke, 2017). Therefore, the purpose of this study was to explore the possibilities and feasibilities of online assessment of scientific inquiry skills, examine the psychometric properties of the test (Scientific inquiry skills test) and ascertain the ability level of the 9th and 11th graders, as well as to examine how the background variables (gender, grade, parents' level of education) affect performance.

5.2.2 Methodology

5.2.2.1 Participants

The sample of the study was drawn from a secondary school in the Omusati region, in the northern part of Namibia. The school comprises grades 8 to 12 and accommodates students from nearby villages and the whole northern part of Namibia. It is a multicultural (grammar) school. Participants were the ninth and eleventh graders (N=118, 44 boys and 74 girls; age mean=16.42, and standard deviation=1.25. Forty-one students were ninth graders (17 boys and 24 girls), age mean=15.10; SD=.67. The eleventh grade consisted of 77 students (boys=27 and girls=50, age mean=17.13, SD=.85). The school has a hostel (dormitory), and all the learners were accommodated in the school hostel at the time of the study. Table 18 shows the distribution of sample and their parents' level of education.

Table 18. Characteristics of the parents' educational level

Parents' level of education		Mothers %	Fathers %
1.	Did not go to school	2.5	3.4
2.	Did not finish primary school	3.4	1.7
3.	Primary education	8.5	5.9
4.	Secondary education	47.5	30.5
5.	Higher education	38.1	46.6
6.	PhD degree	-	2.5
7.	I do not know	-	9.3
Total		100	100

5.2.2.2 Instrument

This study was based on a test of scientific inquiry skills, where students were required to use their cognitive skills (scientific reasoning) to answer questions based on different sub constructs of scientific processes. As in the previous studies (Csapó, 1997; Pásztor et al., 2017; Kambeyo & Wu, 2017; Kambeyo & Csapó, 2017), where inductive and scientific reasoning skills test were used, the idea was to measure different thinking and reasoning skills essential for learning science and learning in general.

The Hungarian based scientific inquiry skills test was adapted and used in the Namibian context. The Magyar Tudományos Akademia (MTA-SZTE) - Research Group on the Development of Competencies of the Institute of Education, University of Szeged, developed the test. The test assesses Hungarian students' inquiry skills. Items were developed based on Wenning's (2007) scientific inquiry skills framework model (see Chapter 4 for further details). The online assessment tool for scientific inquiry skills consisted of 38 tasks with 100 items. Figure 13 shows sample tasks from data handling construct and hypothesis formulation or hypothesis identification respectively.

5.2.2.3 Validation of the research instrument

In order to develop valid items for these scales, the author conducted content analysis of the science syllabi for the grades 8-12 curriculum. A thorough study and comparison of the items' content fit into the Namibian science curriculum was undertaken. The test was also sent to the Chief Science Education officer at the curriculum development centre in Namibia, National Institute for Educational Development (NIED). Furthermore, two science education lecturers from the University of Namibia and three experienced teachers in science subjects were asked to check the contents and the questions of the instrument. Their suggestions were positive and useful

suggestion to improve the language was provided. Tasks that were context/cultural embedded were replaced with the non-contextual tasks. However, since science is a universal subject there were not many culturally embedded items.

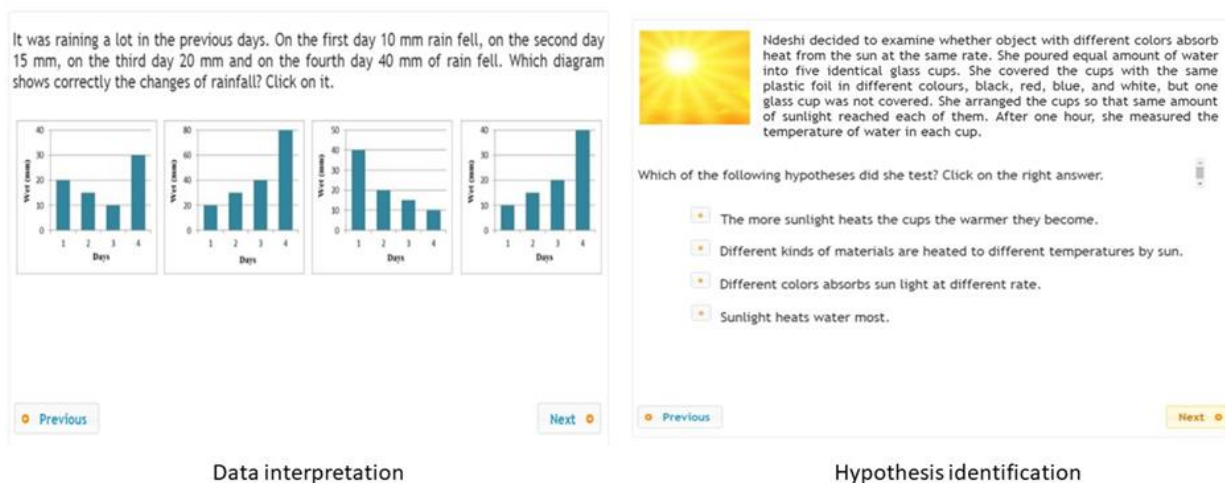


Figure 13 Sample items; Data handling and hypothesis formulation (Nagy et al., 2015)

5.2.2.4 Procedures

The online data collection was carried out through the eDia platform in January 2017, via the internet in the school's information communication technology (ICT) room. Each participant was assigned a number (participants ID) to log into the eDia system, after which the system interface displayed the instructions that students needed to follow. Students used the keyboard, choosing the right answer by clicking or by dragging and dropping figures on the screen with a mouse. The administration of the test took approximately 135 minutes.

5.2.3 Results

5.2.3.1 The psychometric properties and testing applicability

The reliability at the whole scale was high, (Cronbach alpha = .94), see table 18, and two subscales emerged with a better reliability Cronbach alpha of .94 and .85 respectively. However,

the remaining five subscales reliability needs improvement as their reliabilities fluctuated between (Cronbach alpha of $> \text{ or } = .70$). The range of scores on the pilot test was from (< 11 to 100%). The test mean score was (66.70%) with a standard deviation of 15.55% and a standard error of measurement of 1.43% . The test's Cronbach alpha, almost matched the inquiry literacy test by Wenning (2007), when it was first piloted (Cronbach alpha $=.88$), given the fact that this test was self-developed based on the framework by (Korom, 2012). Table 19 shows the descriptive statistics of the test. Students' performance was above (50%), although it could have been better. Students performed above average in subscale data handling and below average in the last three subscales (planning of variables, experimental plans and making conclusions).

Table 19. Descriptive statistics of the whole scale and subscales.

	No. Of items	Cronbach alpha	Mean (%)	SD (%)
Scientific Inquiry Test	99	.94	66.70	15.55
RQs identification	17	.53	73.83	14.23
Hypothesis Identification	13	.56	78.35	15.32
Identify variables	13	.70	58.67	21.61
Planning of variables	11	.94	31.66	37.11
Experimental plans	9	.85	47.83	30.32
Data handling	15	.54	81.69	12.96
Making conclusions	21	.71	74.41	15.44

Note: (RQs= Research questions)

Table 20 presents the correlation matrix showing bivariate relationships between the variables and the whole scale. Moderate strong positive correlations were found among the subscales. Strong correlations between the whole test and subscales was found. However, lower correlations were found between experimental plans and hypothesis formulation, making conclusions and hypothesis formulation than the rest of the scales. Therefore, further improvement on the organizing of items into appropriate subscale is needed in order to yield strong correlations between the subscales and to improve the reliability coefficients at the subscales level. A point to note is that at the whole scale level, there are strong positive correlations with the subscales and the internal consistency coefficient is reliably high.

Table 20. Correlations coefficients between 7 sub scales and the whole test.

	1	2	3	4	5	6	7
1. Whole scale (SI)	1						
2. Data handling	.75	1					
3. Identify variables	.78	.61	1				
4. RQs formulation	.82	.50	.65	1			
5. Hypothesis formulation	.67	.50	.45	.52	1		
6. Planning of variables	.90	.58	.61	.70	.58	1	
7. Experimental plans	.80	.54	.50	.58	.43	.73	1
8. Making conclusions	.80	.58	.51	.60	.42	.64	.60

Notes: all correlations are significant, $p < .01$

The relationships between the dependent variable (Scientific inquiry skills) and the seven subscales was further explored through multiple linear regression analysis. For this analysis we used multiple regression ($F=679.37$, $p < .001$), with an enter method and found that all 7 independent variables contributed 100% to the model. The variables which contributed more to the model are planning of variables 23.8%, making conclusions contributed about 17%, experimental plans and identify variables each contributed about more than 14% respectively. The hypothesis identification was the least contributor see Table 21, while the remaining two variables each explained only above 9% of variance of the scientific skills test.

Table 21. Linear regression analysis for scientific inquiry skills as dependent variable

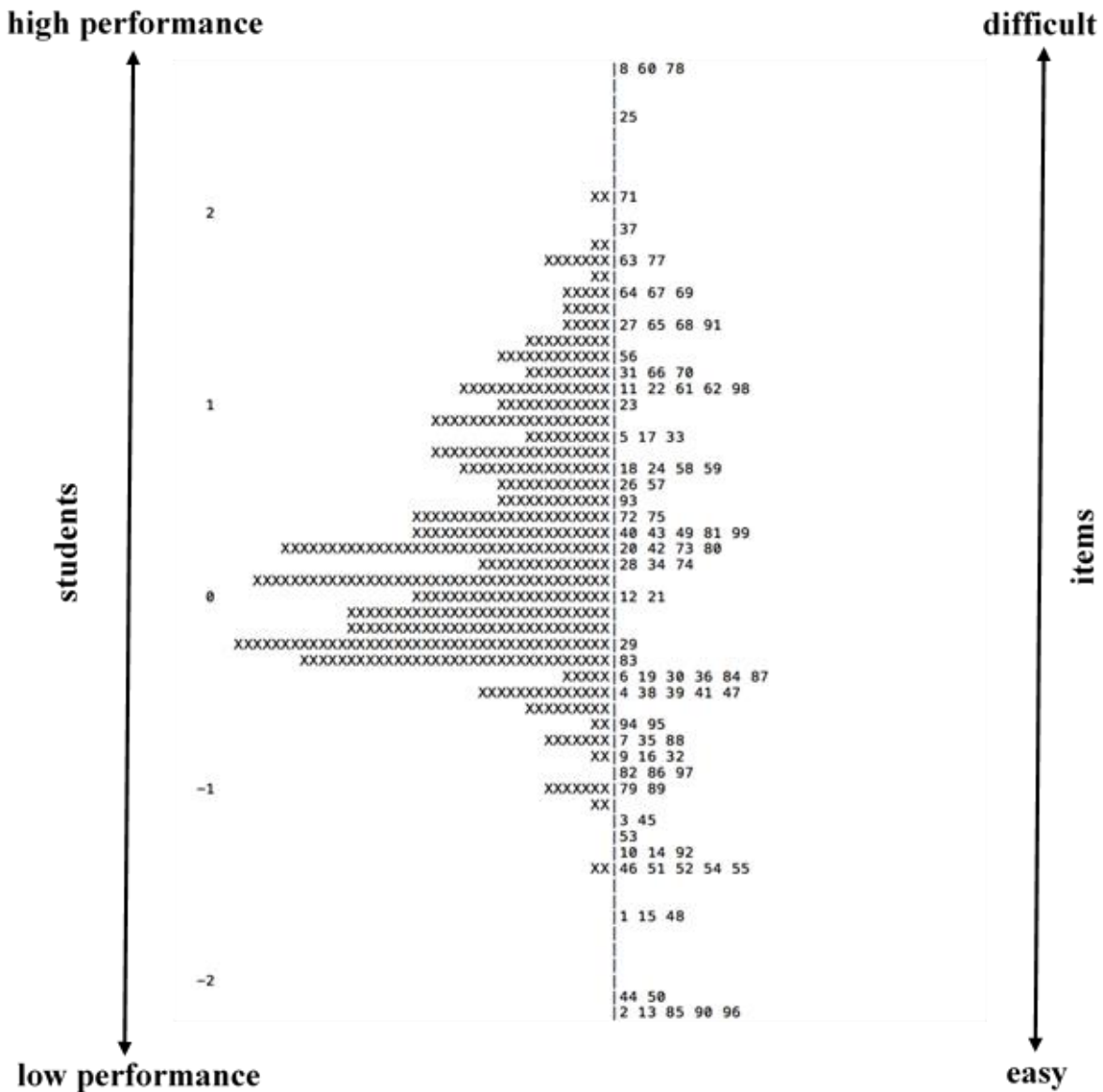
Independent variables	B	SE	t	p	r	Explained variances %
Data handling	.15	.00	-	-	.75	9.44
Identify variables	.13	.00	-	-	.78	14.17
RQs identification	.17	.00	-	-	.82	12.88
Hypothesis Identification	.13	.00	-	-	.67	8.70
Planning of variables	.11	.00	-	-	.90	23.80
Experimental plans	.09	.00	-	-	.80	14.18
Making conclusions	.21	.00	-	-	.80	16.82

Note: Dependent variable: Scientific inquiry skills.

5.2.3.2 Students' ability level

In the person item map (figure 14), the distribution of the test was quite good. The difficulty level of the test fitted the students' ability. Students of different ability were fairly discriminated against by the test which shows that the test was an appropriate measure for the cohorts of students in the region. However, further study needs to be carried out with a large sample for the results to be more generalizable. The distribution is normal, and more students could answer the test items with a probability of more than 50%. Fewer students answered all the items and scored 100%. Item 60, 76 and 78 seemed to prove difficult for the students (see figure 14, figure 15 & figure 16), so

a thorough review of these items is required. Therefore, most students' ability level proved to be above average. More items are needed in order to better and fairly distinguish the ability level of students.



Each 'X' represents 0.9 cases

Figure 14 The person item map of the scientific inquiry skills test

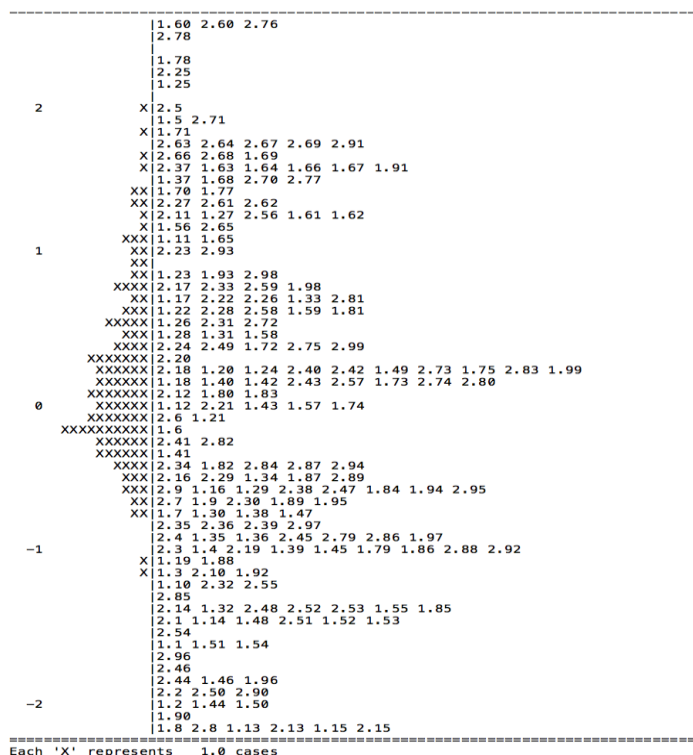
5.2.3.3 Grades and gender differences

No significant differences were found between grades and genders in their performances (see table 22 & 23 as well as figure 15 & 16). The internal consistency in each grade proved to be highly reliable in terms of the whole scale. This may mean that the teaching and learning of science at 11th grade focuses more on examinations than on enhancing reasoning skills and incorporating inquiry methods. Of late, the Namibian education system has become examination-oriented, where passing examinations especially in the externally examined grades (10th & 12th grades) is the only benchmark for performance because there is hardly any monitoring of learning achievements at other levels within the education cycle (Simasiku, Kasanda, & Smit, 2015).

Table 22. The grades mean difference

Grade 9 (N=41)			Grade 11 (N=77)		
Cronbach's α	Mean (%)	SD (%)	Cronbach's α	Mean(%)	SD (%)
.94	67.33	15.35	.94	66.36	15.74

$$t(116) = .32, > .05$$



Note: 1=grade 9; 2=grade11

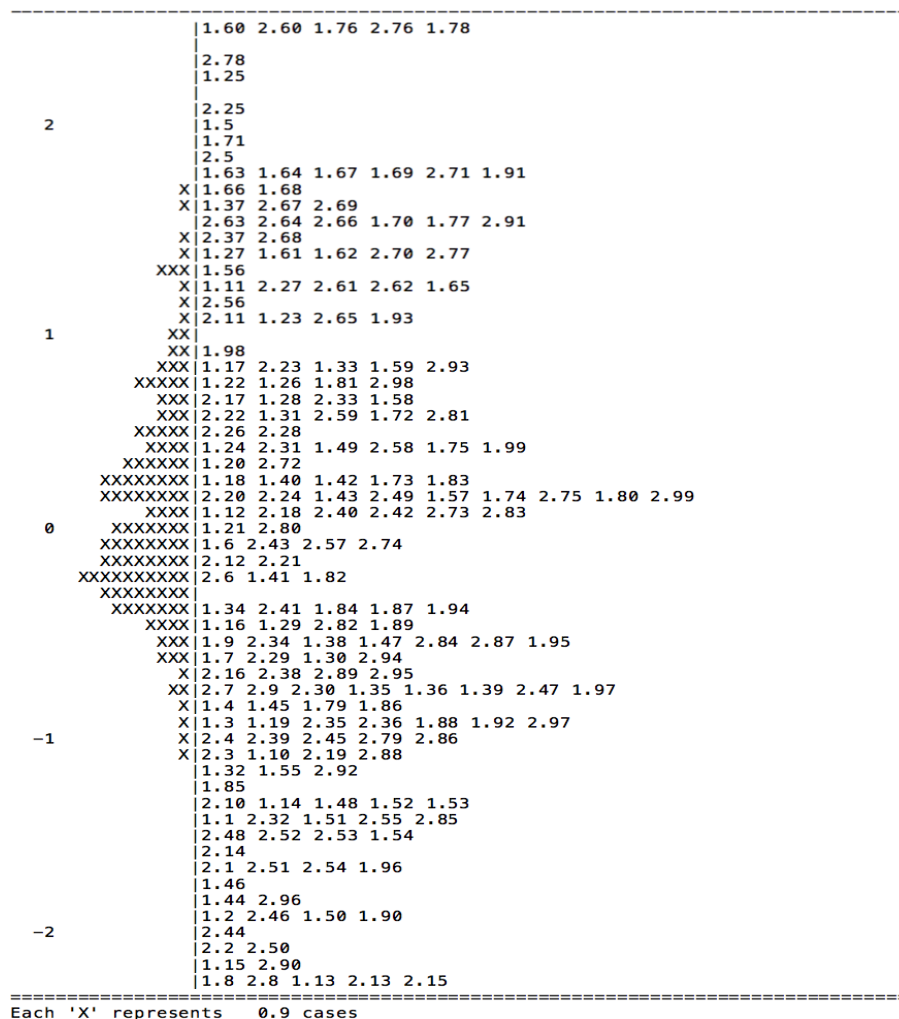
Figure 15 Rasch analysis of the grades performance on each item

The Rasch analyses indicate that the test was not too difficult for both male and female students. Most students found the items very easy. However, items 25, 60, 76 and 78 (see figure 16), proved to be very difficult for both genders. In the same vein, item 8, 13 and 15 were too easy for both genders. No significant gender differences in performance were found.

Table 23. The genders mean difference

Boys (N=44)		Girls (N=74)	
Mean (%)	SD (%)	Mean (%)	SD (%)
65.47	13.35	67.43	16.77

t (106) = .49, p > .05



Note: 1=males; 2=females

Figure 16 Rasch analysis of the gender performance on each item

5.2.3.4 Parents' level of education on students' performance

One-way ANOVA was used to investigate whether parents' level of education has a significant influence on the children's performance. There was no statistically significant difference found between students whose mothers attained secondary education level and higher, and those whose mothers did not go to school, ($p > .05$). However, students whose mothers did not finish primary school, performed significantly better than those whose mothers have reached secondary education and higher education level, ($F = 4.19$, $p < .05$). Fathers have no significant influences on students' performances. Rasch model analysis was also used to indicate the effect of mothers' level of education on students' performance in each item (see table 19 & figure 15). As it was revealed by ANOVA, Rasch analyses also confirmed that students whose mothers did not go to school found the items very easy and they scored with the probability of 50%, (see figure 15). Among the items that proved to be difficult, no students from category 1 to 4 (i.e. students whose mother did not go to school up to students whose mother reached only secondary school level) were found, but only among those whose mothers reached a higher level of education.

		4.60	5.60	3.76	4.76	1.78	4.78	5.78
		4.25	5.25					
		2.60	3.78					
		1.63						
2	X	4.5	5.5	3.25	4.63	5.63	4.64	5.64
		4.71	5.71					
		4.68	5.68	1.91				
	X	4.66	5.66	4.67	5.67	4.91	5.91	
	X	3.5	3.63	3.64	3.69	1.77		
	X	4.70	5.70	3.71	4.77	5.77		
		1.27	4.37	5.37	1.62	3.68		
	X	4.27	5.27	4.62	5.62	3.66	3.67	3.91
	XX	1.11	4.56	5.56	4.61	5.61	4.65	5.65
	XX	4.11	5.11	3.70	3.77			
	X	1.23	3.37					
	XX	3.27	3.62	2.66	2.67			
	XX	4.23	5.23	3.56	3.61	3.65	4.93	5.93
	X	3.11	2.70	2.77	1.98			
1	XX	4.17	5.17	1.22	1.59	4.98	5.98	
	XXX	4.22	5.22	1.26	2.27	4.33	5.33	4.59
	XX	3.23	4.26	5.26	1.28	1.31	2.56	1.58
	XXX	4.28	5.28	4.31	5.31	4.58	5.58	1.72
	XXX	3.17	4.72	5.72	1.75	3.98		
	XXXX	1.18	1.20	3.22	3.33	1.49	3.59	4.75
	XXX	4.20	5.20	2.23	3.26	1.42	4.49	5.49
	XXXXXXX	4.18	5.18	4.24	5.24	3.28	3.31	4.42
	XXXXXX	1.12	2.17	1.40	1.43	1.57	3.72	1.74
	XXXXXX	4.12	5.12	1.21	2.22	2.33	4.40	5.40
	XXXXXXX	4.21	5.21	2.26	3.49	4.74	5.74	3.75
	XXXXXXX	4.6	5.6	3.18	3.20	3.24	2.28	2.31
0	XXXXXXX	2.72	3.83					
	XXXXXXX	3.12	3.40	4.41	5.41	3.43	3.57	3.74
	XXXXXX	3.21	2.49	4.82	5.82	2.99		
	XXXXXXX	3.6	2.20	2.24	1.29	1.34	1.84	4.87
	XX	1.16	4.34	5.34	2.83	4.84	5.84	4.94
	XX	1.9	2.12	4.16	5.16	4.29	5.29	1.38
	XX	4.7	5.7	4.9	5.9	1.47	3.82	4.89
	XX	4.38	5.38	4.47	5.47	3.87		
	XX	1.19	1.30	3.34	1.35	4.36	5.36	3.84
	X	1.4	3.16	4.19	5.19	3.29	4.30	5.30
		3.7	3.9	4.39	5.39	4.45	5.45	4.79
		4.3	5.3	4.4	5.4	3.38	3.47	2.82
	X	4.10	5.10	2.34	3.36	2.84		
		1.14	2.16	3.30	3.35			
		2.9	3.39	3.45	4.55	5.55	3.79	3.86
		3.3	3.4	4.14	5.14	4.32	5.32	2.38
		3.10	4.48	5.48	4.52	5.52	1.53	
		4.53	5.53					
		4.1	5.1	2.39	2.45	4.51	5.51	1.54
		2.4	3.32	4.54	5.54	3.85	2.86	
		3.48	3.52	1.96				
		1.46	3.53	4.96	5.96			
		3.1	4.44	5.44	4.46	5.46	3.51	
		4.2	5.2	2.32	1.50	3.54		
		4.15	5.15	2.48	4.50	5.50	2.53	4.90
		5.8	1.10	2.10	1.13	2.13	3.13	4.13
		1.1	2.1	1.2	2.2	3.2	1.3	2.3
-2		1.1	2.1	1.2	2.2	3.2	1.3	2.3
		1.6	2.6	1.7	2.7	1.8	2.8	3.8
		4.8						

Each 'X' represents 1.0 cases

Note: 1-3=low SES; 4-6=high SES

Figure 17 Effect of mothers' education level on students' performance on items

5.2.4 Discussions

The online assessment instrument for scientific inquiry skills proved to be reliable regarding the whole test. However, at the subscales level, improvements are needed specially the three subscales (data handling, RQs formulation & hypothesis formulation) with Cronbach alpha of $< .65$, see table 19. The remaining four subscales have quite good internal consistency Cronbach alpha of $> .70$.

Research has shown that the advantages of technology-based assessment, such as online test administration and automatic calculation of scoring, reduced the time and cost of the testing process (Pásztor, Molnár & Csapó, 2015). Considering these characteristics, the first steps to be undertaken is to make the instrument suitable for everyday school practice and for possible large-scale assessments in Namibia. The findings indicate that online assessment may provide teachers with an easy-to-use instrument for monitoring the development of students' scientific inquiry skills and reasoning skills and may contribute to the development of effective teaching and learning methods. The question yet to be answered is: are the stakeholders in the education fraternity ready to improve the ICT infrastructures in most public schools as per their blueprint? (MoE, 2001-2006). From my personal experience of being a teacher in public schools for quite some time, it is clear that not all public schools have reliable internet connectivity and functional ICT rooms. In schools that do have ICT rooms, the internet connection may be either weak or there is no signals at all.

The correlations among the subscales proved to be strong and positively significant. This means that with further development, the test could be used to effectively to assess the abilities of students' scientific inquiry skills. Furthermore, students' mean performance was moderate but it could be improved with further training in scientific inquiry skills during their teaching and learning. One question that arose was: does teaching and learning science in Namibia involve the enhancement and inculcation of the necessary inquiry skills for students as it is prescribed in the Namibian national syllabi for natural sciences subjects? Mean performances between the two grades and genders was not statistically significant different.

With regard to genders differences, some science education researchers have reported that gender influenced students' understanding and their attitudes toward science (Al-Zoubi, El-shar'a, & Al-Salam, 2009; Dimitrov, 1999; Lappan, 2000; Valamides, 1996). The study results conformed to other research that no significant difference was found in performance between genders. Recent

research by (Piraksa, Srisawasdi, & Koul, 2014) indicated that gender does not significantly influence students' scientific reasoning ability. Therefore, the findings point to the fact that there is a critical area for improvement of students' scientific inquiry skills and reasoning ability. This also implied that instructional pedagogy in science classroom should put greater emphasis on how to: (i) reason casually based on hypothesis generation, and (ii) design well thought out science experiments, in order to enhance the development of students' scientific inquiry skills and reasoning ability.

There is consistent evidence that parents' education predicts children's educational outcomes, alongside other family characteristics such as family income, parents' occupations, and residence location (Eccles & Davis-Kean, 2005). Other current research echoed these sentiments, e.g. significant positive relationship between parents' education level and academic achievements of students in India (Asad khan, Iqbal, & Tasneem, 2015). Unexpectedly, the results indicate the opposite. Students whose mothers did not finish primary education outperformed those students whose mothers had attained secondary and higher education levels. One explanation for this is that in Namibia, children from low-income brackets tend to work harder than those from affluent families because they want to break of the poverty cycle and live a better life. On the other hand, children of parents with better education do not seem to see the need to work hard, since they are assured of having what they want. However, this is merely an untested idea that needs to be proven with scientific research in order to shed more light on the matter.

So far, the reliability and validity of the scientific inquiry skills have been established. The next step is to assess all the three skills together on one sample with the aim of ascertaining the ability level of the Namibian students in the requisite skills of the 21st century. Sub study 3 was carried out using the LCTSR test (Lawson, 1978, 2000), SI test (Korom et al., 2017) and IR test (Pásztor et al., 2017).

5.3 Sub study 3. The relationships and assessment of scientific reasoning, scientific inquiry and inductive reasoning with students' socio-economic status.

5.3.1 Introduction

This sub study 3 was necessitated by the results of the two preceding sub studies where online instruments were used to assess students' reasoning and inquiry skills. The results of the two studies suggested that online assessment might not be a viable option in Namibia for the next few years or so, due to the poor ICT and internet infrastructures at most of the public schools in Namibia. Therefore, this study presents the results of the paper and pencil assessment of scientific inquiry, scientific reasoning skills and general reasoning skills such as inductive reasoning skills, needed by the students in the 21st century. In the past, education was about instruction but the landscape has changed and students now need to develop expertise in dealing with a technologically driven society (Csapó & Funke, 2017). Therefore, the aims of the study are to assess the 10th and 12th grade students' abilities in scientific reasoning, inductive reasoning and scientific inquiry skills in Namibia, as well as to examine how the background variables (gender, grade, parents' level of education) affect their performances in these three domains.

5.3.2 Methodology

5.3.2.1 Participants

The type of domains studied here do require composing nationally representative samples; however, due to a limited time frame and the resources at my disposal, creating a national sample was not possible. However, samples that were large enough to bear the major typical characteristics of schooling in Namibia were necessary. Thus, samples were drawn from schools in Oshana, Omusati and Oshikoto region in the northern part of Namibia where about 50% of the population resides. Whole school classes were chosen for group testing, and we tried to achieve the best representation of schools in the area in terms of quality and type of schooling. "Other surveys on school achievement have indicated that the target area of our study did not differ much from the national norms in terms of school achievements" (Csapó, 1997, p. 613).

In order to achieve a representative sampling, the samples were chosen from larger secondary schools (8th to 12th grades schools). The students who participated in the study were in

the 10th & 12th graders. The youngest age when the tests can be used is the 7th grade of primary school as below this age reading difficulties are anticipated (Csapó, 1997). The oldest age group that can be tested within the educational system is the 11th graders and beyond. Between these two points, measurements took place. Further data on the samples are summarized in table 24.

Table 24. The samples of the study

Grades	N	Boy/girl %	Mean age	SD
10	130	33.8/66.2	15.68	.68
12	452	36.9/63.1	17.84	.87

Another desired result was to find out if the background variable such as the socio-economic status of the students influences their achievement in the three domains tested. One of the socio-economic indicators is the parents' level of education. Table 25, shows the samples' parents' level of education.

Table 25. The distribution of the level of education of the students' parents for sub study 3

Educational level	Mothers (%)	Fathers (%)
1. Did not go to school	12.9	15.5
2. Did not finish primary school	5.8	4.0
3. Primary education	14.6	16.0
4. Secondary education	38.3	34.5
5. Bachelor degree	18.7	16.0
6. Master's degree	7.2	8.9
7. PhD degree	1.5	1.9
8. I do not know	0.9	3.3
Total	100	100

5.3.2.2 Instruments

The Lawson classroom test of scientific reasoning skills (LCTSR), scientific inquiry (SI) and inductive reasoning (IR) were used for this study where students are required to use their cognitive skills (scientific reasoning and general thinking skills inductive reasoning skills) to answer questions based on different sub-constructs of each domains. The tests assess different thinking and reasoning skills essential for learning science and learning in general. The background variable such students' socio-economic status was also used.

The Magyar Tudományos Akademia (MTA-SZTE) - Research Group on the Development of Competencies of the Institute of Education, University of Szeged, develops these tests, scientific reasoning; scientific inquiry skills and inductive reasoning (see details in Chapter 4). These tests were adapted and used in the Namibian context.

The tests were used in the previous two studies using the online platform and the results were not encouraging. The participants were 5th and 7th graders, much younger than the current samples. Scientific reasoning had a reliability of (Cronbach alpha = .70), but at the subscales level the reliability was so low that it could not be interpreted (Kambeyo, Pásztor, Korom, Nemeth & Csapó, 2017). Therefore, this test was replaced by the LCTSR. The inductive reasoning's reliability index was good (Cronbach alpha = .85) and all the four subscales yielded better reliability from (Cronbach alpha = 74-82). Research shows that the inductive reasoning is not affected by the media (Csapó, 1997; Kambeyo & Csapó, 2017). The scientific inquiry test was piloted on the younger group (9th and 11th grades). The internal consistency was good at (Cronbach alpha = .87). However, there were two subtests with low reliability but the remaining five subtests were within the range of acceptable index of (Cronbach alpha > .70). Table 26, 27 and 28, show the descriptive statistics of the samples on the three tests. The statistics revealed improved reliability indices as compared to when it was piloted.

Table 26. Whole sample descriptive statistics for LCTSR

Tests/subtests (N=582)	No of items	Cronbach Alpha	Mean (%)	SD (%)
LCTSR	24	.90	61.95	26.58
Conservation of volume & mass	4	.72	68.00	34.02
Proportional reasoning	4	.88	49.53	42.55
Control of variables	6	.82	61.54	34.70
Probabilistic reasoning	6	.75	70.22	30.51
Hypo-deductive reasoning	4	.60	37.71	22.03

Table 27. Whole sample descriptive statistics for scientific inquiry skills

Tests/subtests (N=582)	No of items	Cronbach Alpha	Mean (%)	SD (%)
Scientific inquiry	100	.89	72.65	12.68
Data handling	15	.71	90.01	14.93
Variables identification	13	.80	64.30	27.21
RQs identification	17	.60	67.23	15.61
Hypothesis identification	13	.54	75.47	16.63
Variables planning	11	.72	77.55	20.83
Planning experiments	10	.84	64.74	29.44
Making conclusions	21	.70	69.25	15.84

Table 28. Whole sample descriptive statistics for Inductive reasoning skills

Tests/subtests (N=582)	No of items	Cronbach Alpha	Mean (%)	SD (%)
Inductive reasoning	66	.93	77.00	17.00
Figure series	15	.82	80.54	19.32
Figure analogy	15	.81	83.23	18.94
Number analogy	14	.81	66.95	23.58
Number series	22	.91	75.77	24.42

5.3.2.3 Procedures

Informed by the two previous online pilot studies, where students were transported from their schools to the University of Namibia's ICT labs the results showed that the tests were problematic for the 5th and 7th graders. We therefore decided to move to a much older sample, i.e. to the secondary school phase. The data collection was done through paper and pencil method due to the unreliable internet connections in the schools. Students were given around two hours to complete each test. The researchers with the help of the school teachers scored the test. The procedures of group testing were applied in each case.

The data collection took place in the winter (June/July) of 2017. Paper and pencil methods was used. With the help of some teachers at school, I administered the tests after the normal school hours to students. I decided to stay with the secondary school phase and focused on the two grades that write the external examination in Namibia, 10th and 12th grades, also the tests could be useful for them in the near future.

5.3.3 Results

5.3.3.1 Differences between the age groups in all three domains

In the first step, the tests results were computed as percentages of the maximum score. The means and standard deviations of the percentage scales are summarized in Table 29, 30 and 31. The means indicate the percentage of students giving the correct answers. As a general tendency, larger standard deviations were found in the younger age groups. An analysis of the frequency distributions showed that this larger variability could be attributed mainly to a few students in the younger age groups who did much better than their peers.

Table 29. Grade 10 and 12 descriptive statistics of LCTSR

Tests/subtests	Grade 10 (N= 130)				Grade 12 (N=452)		
	No of items	α	M (%)	SD (%)	α	M (%)	SD (%)
LCTSR	24	.89	60.93	25.75	.91	62.24	26.84
Conservation of volume & mass	4	.64	66.92	32.08	.75	68.25	34.58
Proportional reasoning	4	.86	53.07	41.88	.88	48.51	42.73
Control of variables	6	.85	65.64	36.00	.81	60.36	34.27
Probabilistic reasoning	6	.80	61.67	34.45	.73	72.68	29.00
Hypo-deductive reasoning	4	.60	36.41	22.33	.60	38.09	22.00

Table 30. Grade 10 and 12 descriptive statistics of scientific inquiry skills

Tests/subtests	Grade 10 (N=130)				Grade 12 (N=452)		
	No. of items	α	M (%)	SD (%)	α	M (%)	SD (%)
Scientific inquiry skills	100	.89	66.89	13.25	.89	74.31	12.03
RQs identification	17	.65	63.94	17.31	.53	68.18	15.00
Hypothesis identification	13	.54	70.83	16.66	.54	76.80	16.40
Identify variables	13	.84	50.95	31.26	.72	68.14	24.67
Variables planning	11	.74	70.27	23.52	.70	79.65	19.52
Experimental plans	10	.87	55.08	31.70	.82	67.52	28.18
Data handling	15	.70	86.87	17.80	.72	90.91	14.00
Making conclusions	21	.70	66.30	16.22	.70	70.10	15.65

Table 31. Grade 10 and 12 descriptive statistics on inductive reasoning

Tests/subtests	Grade 10 (N=130)				Grade 12 (N=452)		
	No. of items	α	M (%)	SD (%)	α	M (%)	SD (%)
Inductive reasoning	66	.93	65.82	18.54	.92	79.80	15.15
Figure series	15	.80	74.31	20.11	.82	82.33	18.73
Figure analogy	15	.84	76.51	22.96	.80	85.16	17.16
Number analogy	14	.70	56.65	18.95	.82	69.91	24.00
Number series	22	.91	58.57	28.02	.88	80.72	20.82

Analyzing the test descriptive statistics further, the statistics revealed that the internal consistencies of the tests were quite good. Analysis of the whole sample showed good performance in all three tests compared to the previous sub studies where online methods were used. This finding is in line with the literature, for example Bunderson, Inouye, and Olsen (1989) state, "...the scores on tests administered on paper were more often higher than on computer-administered tests... the score differences were generally quite small..." (p. 378). Mead and Drasgow (1993) further posit that in a meta-analysis of well-designed computer versus paper-based cognitive ability tests they found that on average, paper-based test scores were very slightly greater than computer-based test scores.

The LCTSR reliability index of the whole sample was high, except for hypothetical deductive reasoning subscale of scientific reasoning which was below the minimum required (Cronbach alpha > .70) in the testing field. Further to that, in both age groups, hypothetical deductive reasoning subtest produced a low reliability index of (Cronbach alpha = .60). The

hypothetical-deductive reasoning is thought to be the most complicated ability in the Lawson's test (Han, 2013), and therefore put last, representing the last stage of formal reasoning. Nevertheless, it is the most important core skill in scientific reasoning as hypothesis testing is always the goal of scientific inquiries and applications of scientific methods.

On the LCTSR test, the last four questions were reserved for this skill dimension (see appendix D). These are fairly long questions requiring a whole page of reading and parsing. The two questions in a pair are not structured as answer explanation any longer. In the first pair (questions 21-22), the first question provides a number of experimental designs while the second question gives a set of possible experimental outcomes. The two questions need to be coordinated in order to form a consistent pair of design and outcome that also can be used to test the provided hypothesis. In the second pair of questions (questions 23-24), the narratives of the questions present an experimental setting and two possible hypotheses. The first question asks for a selection of experimental outcomes that would prove the first hypothesis wrong, while the second question asks for experimental outcomes that would negate the second hypothesis. As a result, in order to respond correctly to these questions, students need to have a well-established reading and comprehension capacity as well as information processing skills to parse out the useful information from an abundant collection of co-existing but not relevant features. The frequency analysis showed that students did not do well on these items (see table 29).

With regard to SI, hypotheses identification and research questions identification subscales of the scientific inquiry, also showed low reliability indices. These questions under this sub scales are also set similar to the LCTSR question under hypothetical deductive reasoning. Like in sub study 2, students did not do well also in these subscales. The questions under these subscales demand a good command of English comprehension and students are required to process more information in order to get the correct answer. These two subscales were not improved as per the findings from sub study 2. Two assumptions were made, 1) the test was to be written by an older sample, 2) the mode of taking a test would be the traditional paper and pencil method, so there was no need to improve it. Furthermore, research questions identification and hypothesis formulation only yielded (Cronbach alpha < .60) in both age groups.

It seems that that these items may need further reorganization before they measure the same thing under these constructs. Or the other plausible explanation could be that students in Namibia might not have reached this level of reasoning. However, the IR proved to be a good assessment

instrument with (Cronbach alphas of $> .80$). In all three domains, both age groups yielded (Cronbach alpha $> .86$). Inductive reasoning domains proved to be a highly reliable test in both age groups as it yielded (Cronbach alphas $> .75$) in both subtests.

Further examination of the test scores showed that there were no gender significant differences in performance in these three tests, scientific reasoning ($t = .05$, $p > .05$), scientific inquiry skills ($t = 1.74$, $p > .05$) and inductive reasoning ($t = .63$, $p > .05$). In terms of age, no significant differences were found concerning scientific reasoning ($t = 1.26$, $p > .05$). However, grade 12 students performed significantly better than the grade 10 students in scientific inquiry skills ($t = 5.19$, $p < .01$) and in inductive reasoning test ($t = 8.72$, $p < .01$).

5.3.3.2 Students' ability level

The Rasch analysis for the items in inductive reasoning showed the following values: item fit: $0.82 < \text{weighted mean square (wMNSQ)} < 1.12$; expected a posteriori/plausible value (EAP/PV) = 0.91; and difficulty: $-.343 < r_{it} < 2.244$ (Adams & Wu, 2002). The Wright item map for all the three tests almost showed a good match between item difficulty and the students' abilities see (figure 18). However, in scientific reasoning and inductive reasoning, the tests could not differentiate high ability students. In case of scientific inquiry, there were fewer items that could differentiate the high ability students.

The distributions of all the three tests are normal, and more students could answer the test items with a probability of scoring more than 50%. As indicated in figure 16, some students found the scientific reasoning test items difficult. With regard to SI, several items proved to be easy i.e. item (1, 2 & 8) most students answered these items with the probability of scoring 100%. Overall, few students would probably score the items correctly with over 90% in all three tests. For scientific reasoning, the item fit: $0.82 < \text{weighted mean square (wMNSQ)} < 1.25$; expected a posteriori/plausible value (EAP/PV) = 0.87; and difficulty: $-1.150 < r_{it} < 1.046$. For scientific inquiry skills, the item fit: $0.67 < \text{weighted mean square (wMNSQ)} < 1.14$; expected a posteriori/plausible value (EAP/PV) = 0.83; and difficulty: $-2.627 < r_{it} < 3.525$.

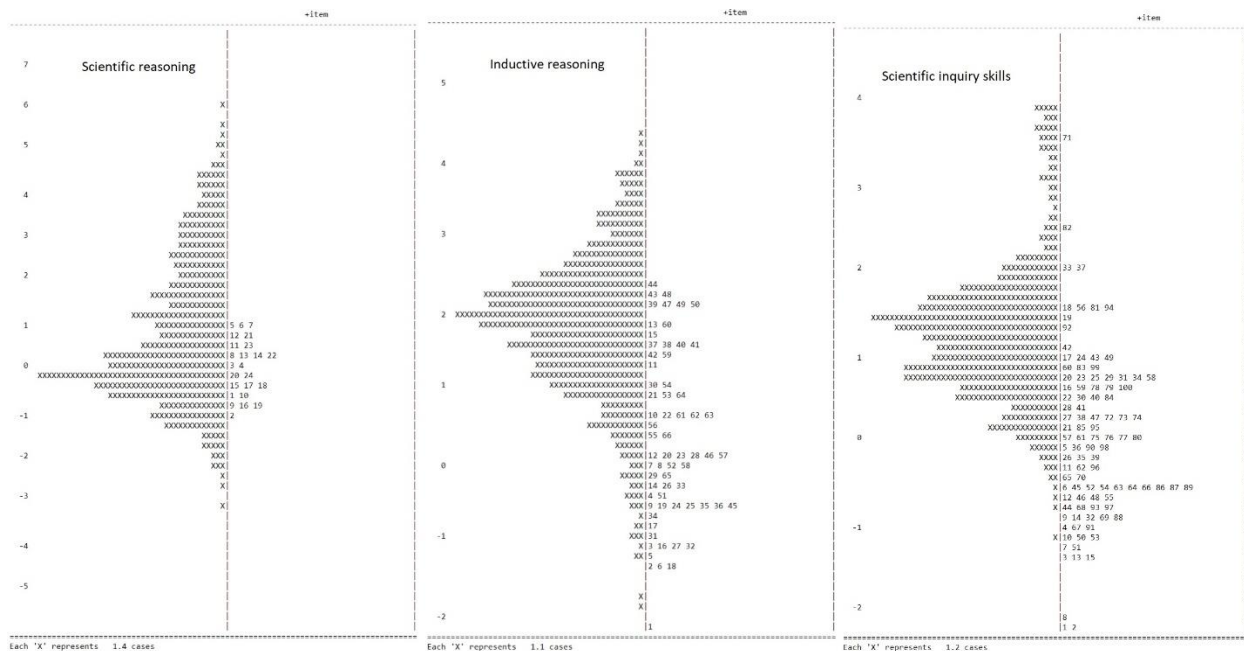


Figure 18 IRT maps for the three domains

5.3.3.3 Analysis by SEM

The simple bivariate correlation was run. The analysis shows unusual results whereby LCTSR and the other two constructs could not significantly correlate, although the literature says otherwise. This leaves rooms for further investigation. However, IR and SI showed significant correlation ($r=.29$, $p<.001$). SEM analysis was further then used to examine the relationships between LCTSR, IR & SI. The SEM model of the three domains that assess domain general reasoning skills (IR) and domain specific reasoning skills i.e. LCTSR and SI showed a good model fit, (see figure 19). IR, LCTSR and SI have been specified as latent factors. Results revealed that four dimensions, figure series (FS), figure analogy (FA), number series (NS) and number analogy (NA) explained inductive reasoning (IR) performance ($.57 < \beta < .79$). Five dimensions explained the performance of scientific reasoning (LCTSR) these are conservation of mass and volume (CON), proportional reasoning (PROP), control of variables (COV), probabilistic reasoning (PROB) and hypothetical deductive reasoning (HYP) ($.60 < \beta < .82$). The third domain, scientific inquiry skills (SI) performance has been explained by seven dimensions which are, data analysis (DA), variable identification (VI), research questions identification (RQs) hypothesis identification (HF), variable planning (VP), planning experiments, (PE) and making conclusions (MC) ($.49 < \beta < .64$). Significant correlation was found between SI and IR ($r=.35$, $p<.01$).

However, no significant correlation was found between IR and LCTSR as well as between SI and LCTSR ($p < .05$).

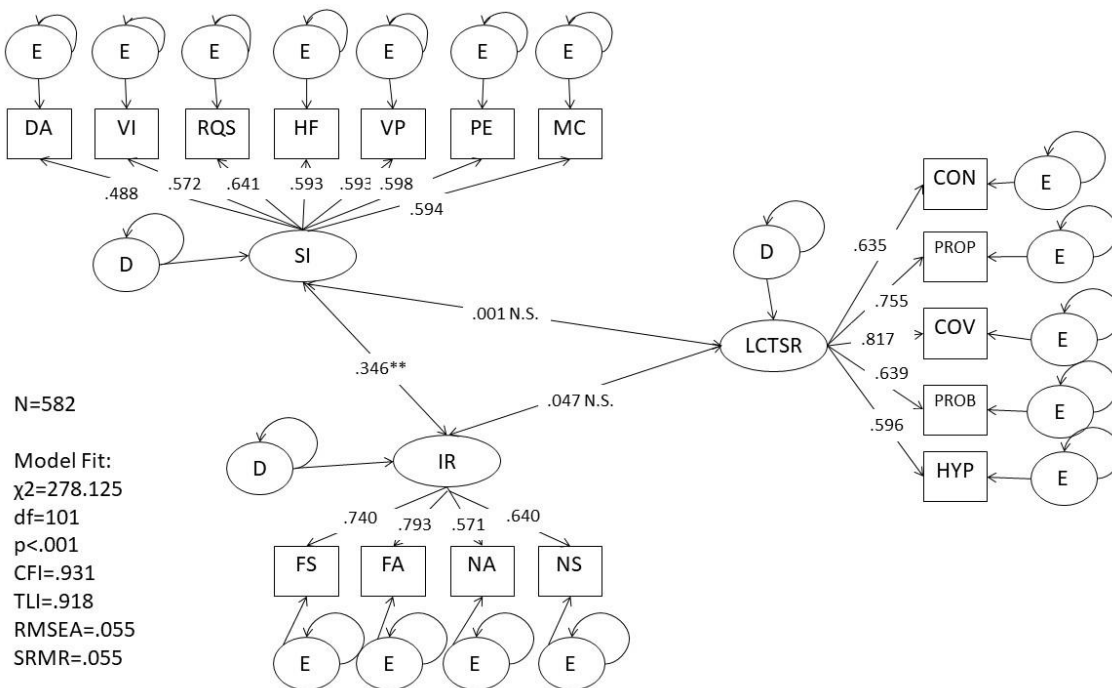


Figure 19 A model of SR, SI and IR with dimensions explaining each domain.

Moreover, the analysis was also carried out per age group of the sample in order to understand more detail on the relationship between the three domains. The model fits were acceptable for both age groups (see figure 20 & 21). However, the 12th grade (older age group) model fits were better than the 10th grade students (younger age group). The same dimensions as in the whole sample model explained the performance of IR, LCTSR, and SI respectively. The 10th grade analysis showed that DA, VP, VI and RQS strongly contributed towards SI performance ($.60 < \beta < .64$). The CON dimension contributed less towards LCTSR ($r = .42$) while the rest of the dimensions contributed more ($.56 < \beta < .80$) respectively. With regard to IR in the 10th grade, the NS and FA dimensions contributed a bit more than the FS and NA, (see figure 20). Significant correlation was found again between SI and IR ($r = .36$, $p < .01$), again no significant correlations were found between SI and LCTSR, LCTSR and IR.

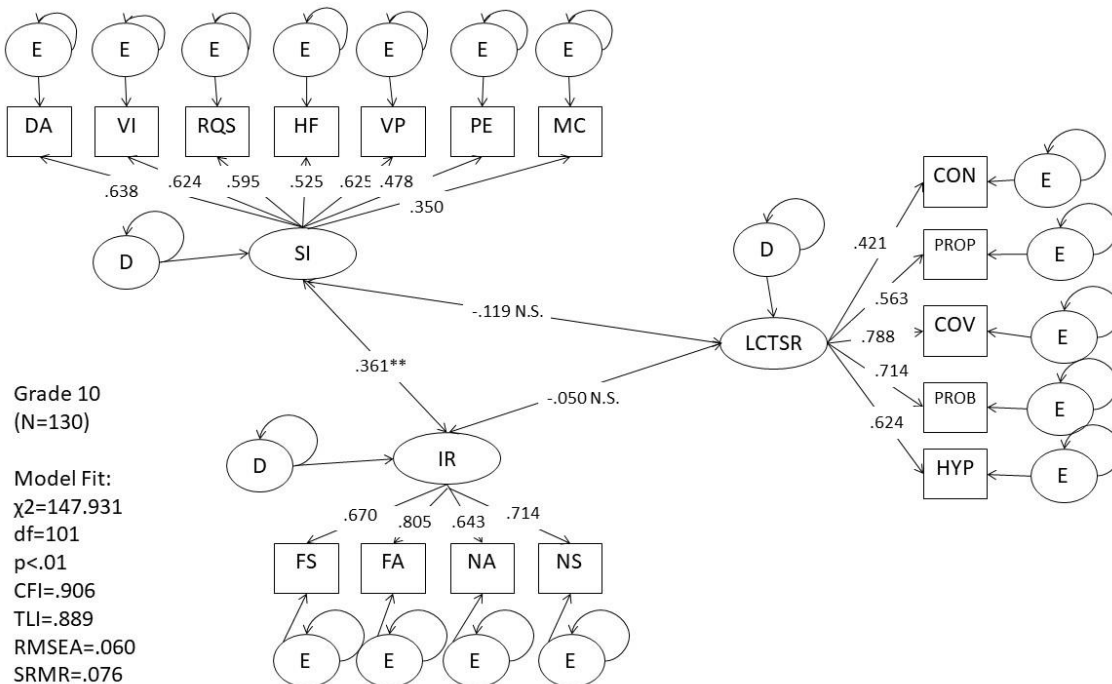


Figure 20 A grade 10 model of SR, SI and IR with dimensions explaining each domain.

The 12th grade analysis showed that, PE, HF, RQS and MC contributed more than the remaining three dimensions towards SI performance ($r = .61-.67$). The strongest contributors towards LCTSR in the 12th grade were the PROP and COV with ($r = .79$ and $.84$) respectively. The remaining three dimensions contributed almost equally with ($r > .60$). With regard to IR in the older age group, the FS and FA dimensions contributed equally with ($r > .75$) while the NS and NA contribution were also in the same range of ($r > .51$) (see figure 21). It is interesting to note that once again, significant correlation was found between SI and IR ($r = .27$, $p < .01$), and no significant correlations were found between SI and LCTSR, LCTSR and IR. One can deduce from this that LCTSR is not significantly related to the other two domains, which is an unusual phenomenon because literature suggests that IR can correlate well with most of the domain specific constructs. These present results may be due to data collection process, the process might have been too long and students lack motivation to taking the test along the way. This presents also an opportunity for further research in this area in Namibia again,

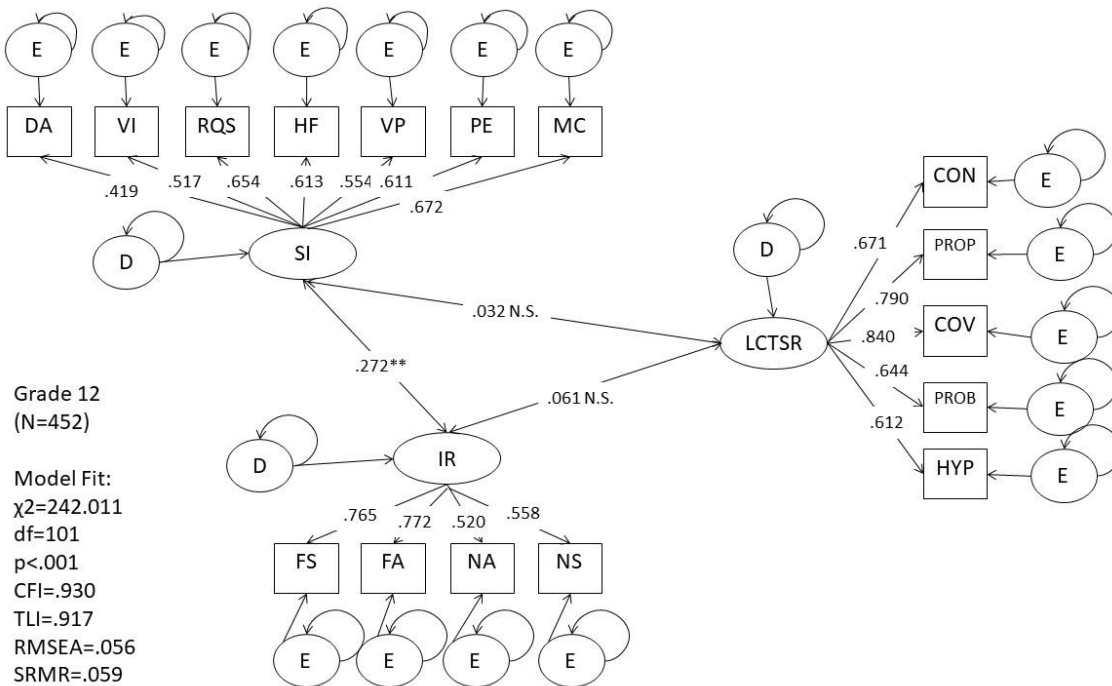


Figure 21 A grade 12 model of SR, SI and IR with dimensions explaining each domain.

Furthermore, analysis of variance (ANOVA) was also carried out in order to find out if parents' level of education has any influence on the students' performance in the three domains (LCTSR, SI & IR). Based on literature, students from good socio-economic backgrounds tend to perform better in school subjects than those from poor socio-economic backgrounds (Keller et al., 2017). ANOVA showed no significant influence from the parents' level of education concerning SR and SI performances. However, significant difference was found between these groups; (*mothers who did not finish primary school and mothers who finished only primary school*), in the case of IR ($p<.05$). Students whose mothers did not finish primary school performed better than those whose mothers went up to secondary school. This is contrary to what is stated in the literature. Fathers' level of education did not significantly impact on students' performance in all three domains at the whole sample level.

We also tried to analyze the data at the age group level. Parents' level of education did not significantly influence the grade 10s students' performance. In the grade 12 group, parents' level of education did not affect students' achievement in SR and SI. However, significant difference was found between the following two groups, (*fathers who did not go to school and fathers who attained Master's degree*) in IR ($p<.05$). Again, students whose fathers did not go to school performed significantly better than those whose fathers went up to a master's degree level. This

finding can be attributed to a difference in cultural background. As was mentioned earlier most of the literature is contextualized in a western setting and there is no Namibian research data on this issue. In Oshiwambo culture, (*a tribe where most of the samples are from*), the non-educated parents encourage and motivates their children to work hard. Although this is not a scientific statement, the results tend to agree with it. Even in the previous study (Kambeyo, 2017), students whose parents did not go to school outperformed the students whose parents have at least above secondary school education.

5.3.4 Discussions

The paper and pencil assessment proved to be very reliable as all the three tests yielded high internal consistency of (Cronbach alpha > .87) in both age group. With regard to scientific reasoning, the original Lawson Classroom Test of Scientific reasoning (LCTSR) was used (Lawson, 1978, 2000). The reliability results conform and surpass the results of Lawson (Cronbach alpha = .78, .77, & .80) respectively in the 9th, 10th, and 11th grade in the USA. With regard to IR, the results match with what is found in the research community within this field ($\alpha > .80$). Consequently, the SI test also proved to be highly reliable. The results matched with what the proponents of scientific inquiry skills found (Wenning, 2007). However, among the seven subscales of scientific inquiry skills tests, two subscales (RQs identification & hypothesis identification) yielded a low reliability of (Cronbach alpha < .70), which is the acceptable Cronbach alpha in the testing field. The reliability indices of these two subscales improved compared to what was found in sub study 2, where online testing was used. This suggests that the items in the two subscales may not necessarily measure the same construct. Therefore, a rearrangement of items is necessary.

Mean performance results indicated that no significant difference was found between genders in all three tests. Concerning age groups, no significant differences were found concerning Lawson classroom test of scientific reasoning. However, the 12th grade students performed much better than the 10th grade students in scientific inquiry skills and in inductive reasoning test. The Rasch analysis of students' ability level showed that all the three tests proved to be a good match between item difficulty and the students' abilities.

SEM analysis revealed the relationship between LCTSR, IR & SI. The SEM model for all three tests showed a good model fit for the whole sample, and for the two-different age groups. As

in the literature, a positive correlation between IR and SI was found (Korom et al., 2017). However, no significant correlation could be found between IR and LCTSR as well as between LCTSR and SI. This is a bit surprising as the literature (Pásztor et al., 2017; Korom et al., 2017, Csapó, 1997; Molnár et al., 2015) indicated that domain general thinking skills correlate and predict most of the domain specific thinking skills. In our previous test where the online assessment method was used, a positive correlation between IR and SR was found (Kambeyo & Csapó, 2017). These unusual results could be caused by data collection process or handling. Further research on this is really advocated for so as to establish the root cause of this.

Again, in this study, parents' educational level predicted students' achievement in the opposite way from the findings in India where a significant positive relationship between parents' education level and academic achievements of students was noted (Asad khan et al., 2015). Students whose mothers did not finish primary education performed significantly better than those whose mothers have attained secondary and higher education levels. One explanation for this is that, in the Oshiwambo culture: *(an ethnic group in Namibia where the sample for the study was drawn from, that constitutes 51% of the Namibian population)*, children from low-income groups tend to work harder than those from high income families.

5.4 Sub study 4. The relationships between 8th graders' second language readings skills, inductive reasoning, scientific reasoning and socio-economic status in Namibia.

5.4.1 Introduction

Subsequently, the results and findings of sub study 1, 2 and 3, raise the question of whether language (English) had an impact on the students' performance since they are assessed in the language other than their mother tongue as compared to large scale assessment programs such as PISA, TIMSS and others, where students take the tests in their home language. A decision was then made to carry out a supplementary study on the relation between reading comprehension and reasoning skills (sub study 4). The samples used in sub study 4 were from the same schools as the samples in sub study 1, 2, and 3, however, this time the 8th grades took part since the reading comprehension test was designed for the 14 to 15 years old (Nikolov & Csapó, 2018). In addition, we also wanted to ascertain the level of the 8th grade students in order to obtain a snapshot of the whole secondary school phase and their ability in LCTSR, and IR skills.

After independence in 1990, English was declared the official language for Namibia. All the official correspondence and teaching in Namibian schools must be in English. The provision was made that the teaching from pre-primary to 3rd grade would be in the mother tongue, then from the 4th grade upwards English must be the language of teaching and learning. Given this situation, and the fact that a significant portion of the population in Namibia lives in rural areas and speak a limited amount of English daily, the following questions came to mind: do the children in Namibia fully comprehend and follow instruction well in English? Could the students' performance be affected by their socio-economic status? These were the some of the after thoughts that resulted in this supplementary sub study. Therefore, using a cross-sectional study like this one, the overall impact of students' SES on English as the Language of Teaching and Learning (LoTL) was explored, and how the possible lack of English comprehension skills could affect students' scientific reasoning and general cognitive abilities (inductive reasoning) achievement.

5.4.2 Methodology

5.4.2.1 Participants

The sample of participants was drawn from the northern Namibia public schools where all the other three sub studies were conducted, representing students in four statistical regions of the country and schools located in settlements of different sizes proportionally. The sample comprised 250, 8th grade students (36% boys) and (64% girls), age mean (M=15.10; SD=.580).

5.4.2.2 Instruments

Four instruments were used. All were validated in previous research (L2 tests in 2002; see Csapó, 2014; Csapó & Nikolov, 2009) both in paper and pencil and online versions prior to use in the present study. The LCTSR test, (Lawson, 2000) and IR test (Pásztor et al., 2017) were used, and both were validated also through online and paper and pencil testing.

- (1) A paper and pencil English reading comprehension test was used to assess students' English reading comprehension skills (table 32 & appendix E);
- (2) Students' scientific reasoning skills were measured by paper and pencil scientific reasoning test (see appendix D).
- (3) Students' general cognitive abilities were measured by paper and pencil inductive reasoning test (see appendix B);
- (4) Participants' socio-educational status was tapped by a short questionnaire (Table 37).

Five English reading comprehension tests were included in the assessments. They were estimated to cover a relatively wide range of proficiency (A1 and lower band of A2 on the *Common European Framework of Reference* levels, Council of Europe, 2001) in line with earlier requirements for 8th graders in the national curriculum.

Table 32. The paper and pencil English reading comprehension tests

Test	Task	Input texts
Reading 1	Match words with appropriate sentence	Definitions of words
Reading 2	Match notices with meaning	Public notices and their meaning
Reading 3	Match questions with answers	Interview from youth magazine
Reading 4	Match questions with answers	Quiz texts for teenagers
Reading 5	Match short advertisements with missing words	Advertisements

The length of the texts in the items ranged from a single word, an expression, or a sentence, to a very short passage. All tests focused on meaning and not form. The rubric was given in English; all tests included an example in order to make sure students knew what to do. The topics and the task types (multiple matching) were familiar to the test students (table 32). Although none of the tests was borrowed from teaching materials, they were similar to the tasks students came across in course books. The vocabulary and structures of the texts were expected to be at or a bit beyond 8th graders' level of proficiency. The number of items in the five tests ranged between eight and ten; the tests comprised a total of 45 items, with different five reading tests assessing the same construct.

The reliability of the RC was high: (Cronbach's alpha .83). This proves that the test can be used in the Namibian context. At the subscale level the reliabilities indices were quite high (see table 33).

Table 33. The reliabilities indices of the reading comprehension tests and its subscales

Scales	Cronbach Alpha
Reading comprehension test	.83
Reading test 1	.70
Reading test 2	.84
Reading test 3	.71
Reading test 4	.77
Reading test 5	.65

The reliability of the LCTSR Lawson's test (Version. 2000) has been evaluated by researchers who used this test. Typical internal consistency in terms of Cronbach alpha's ranged from (Cronbach alpha 0.61 to 0.78) (Lee & She, 2010). The popularly used version of Lawson's Classroom Test of Scientific Reasoning was released in the year 2000. The IR test used in this sub study is adopted from the IR test by (Pásztor et al., 2017). The original test contained four subtests of 66 items, but the one used here consisted of two subtests (i.e. figurative and number reasoning tests) with 29 items, due to time factor.

Students' SES was measured by their parents' level of education on a 7-point scale (see table 37). This questionnaire was used in the previous sub studies and was subsequently improved. Students were asked to specify their mothers' and fathers' highest level of education from these options: (1) did not go to school; (2) did not finish primary school; (3) grade 10; (4) grade 12; (5) bachelor degree; (6) master's degree; (7) PhD degree.

5.4.2.3 Procedures

Data collection for this study was undertaken during January/February 2018. Students were assessed in RC, IR and LCTSR domains by paper and pencil test. The SES questionnaire was also administered by the same method. All instruments were administered during the afternoons after normal school hours. Students had about two hours and 30 minutes to take the inductive reasoning and scientific reasoning tests. The reading comprehension and the SES questionnaire were taken in the morning before lessons started and required about one hour and 45 minutes to complete both tasks.

5.4.3 Results

5.4.3.1 Achievements of students in RC, IR and SR tests and gender differences

As presented in tables 34, 35 and 36, no significant gender difference was found in students' performance in all three tests. The comparison of students' RC, IR and LCTSR ability scores using mean, standard deviation, and independent t-test showed that there were no significant differences in scores between males and females ($p < .05$). These results resonate well with previous research where some educators indicated that no significant difference in the scientific and inductive reasoning between males and females existed. These findings indicate that the two groups were equal in their reading comprehension skills, scientific reasoning and general reasoning skill. The paper and pencil tests proved to be reliable and useful measure for this age cohort. As it was revealed in sub study 3, hypothetical deductive reasoning items still proved to be a challenge to students and may need to be changed as it yielded very low reliability. In the reading comprehension tests, dimension five (subscale 5) proved to be a bit difficult for the children, ($M=66.60\%$) compared to the rest of the dimension (subscales) where students scored more than 70% mean average. Dimension five is mostly concerned with advertisement and filling the gap in the adverts. About IR, students did well in the two dimensions of the used IR tests. Therefore, the paper and pencil method seemed to be a convenient way of assessing these skills, as most of the students are familiar with the medium. As it is indicated in tables 34, 35 and 36, the mean performance of above 80%, indicates that the sample used may be highly motivated as compared to the other sample in study 2 & 3. However, a longitudinal study is needed to determine the ability level of the Namibian students in these skills (IR, LCTSR & RC).

The descriptive analysis of the test was done. The mean average for the three tests were, 78.85% for LCTSR, 88.42% for RC and 84.48% for IR. A paired sample t-test was carried out and the results indicate that the sample performed significantly better in IR than in SR ($t_{(249)} = 6.09$, $p < .05$). The reasons for this result may emanate from the fact that not much reading is required in the IR test, (see appendix B), but reasoning is required in order to derive the rules and get the answer, while with LCTSR, a bit of reading and comprehension of some technical words is required coupled with the reasoning aspect. There was no significance difference in performance between the three tests.

Table 34. Students' achievement in scientific reasoning test

Test/Subtest	Cronbach alpha	No of items	Mean (%)	SD (%)
Conservation	.90	4	82.70	32.98
Proportional	.83	4	76.80	34.23
CoV	.73	6	82.13	24.90
Probabilistic	.75	6	65.60	30.92
Hypothetical	.64	4	92.00	18.09
LCTSR	.74	24	78.85	14.93

Table 35. Students' achievement in reading comprehension test

Test/Subtest	Cronbach alpha	No of items	Mean (%)	SD (%)
D1	.70	10	95.88	10.00
D2	.84	9	94.67	14.74
D3	.71	9	93.07	13.20
D4	.77	8	91.25	17.26
D5	.65	9	66.71	23.41
RC	.83	45	88.42	10.22

Table 36. Students' achievement in inductive reasoning test

Test/Subtest	Cronbach alpha	No of items	Mean (%)	SD (%)
FS	.84	15	87.07	18.17
NA	.80	14	81.46	19.42
IR	.86	29	84.48	16.00

5.4.3.2 Relationships between students' RC and their achievement in IR and SR

Furthermore, the relationships between RC, IR and LCTSR were further explored and compared with one another as well as with findings of previous studies. The strongest relationship ($r=.60$, $p>.05$) is between inductive reasoning and achievement on the English reading comprehension tests. While the relationship between IR and SR, and RC and SR were the same ($r=.48$, $p>.05$) each. The results fit in well with what is found in the current literature where a

strong correlation was found between IR and RC ($r=.57$) where students English was a foreign language (Nikolov & Csapó, 2018).

Paired sample t-test results indicate that the sample performed significantly better in the RC than in both IR and LCTSR ($t=5.05$; $t=11.24$, $p<.05$ respectively). This indicated that the sample did not have difficulties in understanding and following instructions in English as a second language, because of the sample's above average performance in the RC test. This finding further confirmed the results obtained in study 3 and resonates well with what is found in international literature (see Nikolov & Csapó, 2018).

To answer fully the research questions, we also ran a structural equation modelling to check whether the sub dimensions (subscales) of RC predict the achievements of students in the two constructs, IR and LCTSR and these two constructs were both correlated. The model fit for this analysis was not good, (i.e. $\chi^2=108.05$, $DF=51$, $CFI=.88$, $TLI=.85$, $RMSEA=.07$ & $SRMR=.06$). Furthermore, RC on LCTSR and RC on IR was also tried and the model fit was still not acceptable. Then, we tried to build a model whereby RC as a latent variable which is made up by its five dimensions which are predicting IR, the model fit for this was also not acceptable (i.e. $\chi^2=37.84$, $CFI=.92$, $TLI=.87$, $DF=13$, $RMSEA=.09$, & $SRMR=.06$). Then we removed LCTSR from the model and kept only the sub dimensions of RC and IR and the model fit improved, $\chi^2=8.56$, $DF=4$, $CFI=.97$, $TLI=.92$, $RMSEA=.07$ and $SRMR=.03$, (see figure 22).

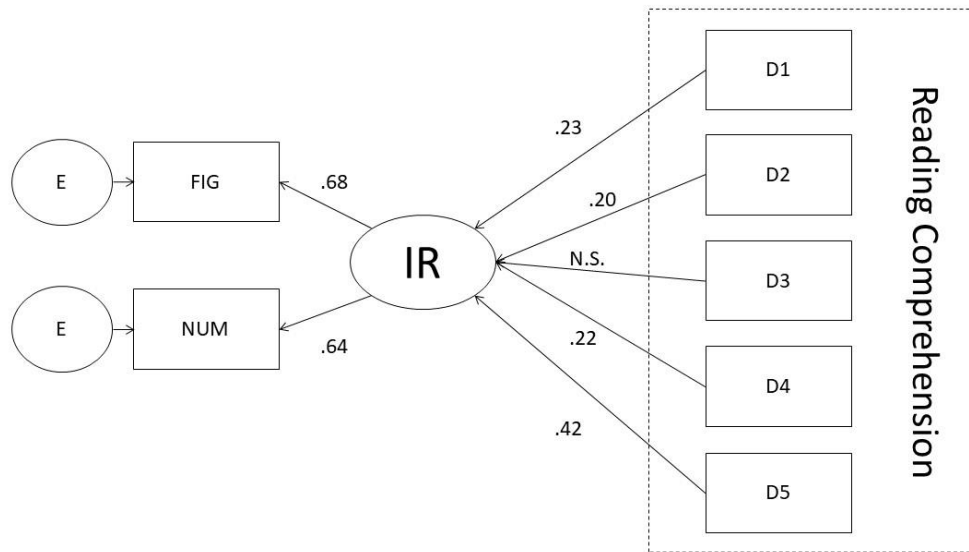


Figure 22 A model of RC, dimensions explaining the achievement of IR.

IR consists of two dimensions in this test, figural and number reasoning, that are highly correlated to IR. RC consists of five dimensions, which are; D1 (definitions of words), D2 (public notices and their meaning), D3 (interview from youth magazine), D4 (quiz texts for teenagers), and D5 (advertisements). Four dimensions of RC significantly predicted the IR. However, one dimension of RC does not have a significant predicting effect on IR.

5.4.3.3 Relationships between students' SES and their achievement in IR and SR

Previous studies conducted in developed countries over a decade earlier found a relationship to think about between students' SES and choice of second language. The more affluent the parents were, the more probable it was that their children would want to study in English (Csapó & Nikolov, 2009; Nikolov & Csapó, 2010; Nikolov & Józsa, 2006).

Table 37. Distribution of parents' level of education (%)

Parents' educational level	Mothers (%)	Fathers (%)
1. Did not go to school	3.6	7.2
2. Did not finish primary school	5.2	5.2
3. Grade 10	9.6	7.2
4. Grade 12	30.8	25.6
5. Bachelor degree	18.4	18.0
6. Master degree	28.0	29.6
7. PhD degree	4.4	7.2
Total	100	100

As the data in table 37 indicates, about over half (50.8%) of the participants' mothers received higher education, that is to say bachelor to PhD degree. On the other hand, about 54.8% of the participants' fathers received higher education. The difference in education between mothers and fathers' level of education could be attributed to the fact that prior to independence girls and boys never received equal treatment when it came to accessing education.

The results indicate that mothers' and fathers' education level correlates negatively ($r = -.210$; & $-.144$; $p < .01$) respectively with LCTSR. As a combined variable of SES, parents' educational level or students' SES does not influence their achievement on IR, ($r = .016$, $p > .05$). However, students' SES negatively correlate with LCTSR, ($r = -.197$, $p < .01$). As expected, IR and LCTSR were highly correlated ($r = .569$, $p < .01$).

These outcomes are in line with the results of study 3 and 4 of this project, however very different from what is found in the current literature. It has to be understood that the literature read in preparation for this study was based on western culture (developed countries) where students' SES highly related to the students' achievement in LCTSR, IR and reading comprehension skills (Csapó & Nikolov, 2009; Nikolov & Csapó, 2010; Nikolov & Csapó, 2018). The possible explanation and assumption that we could make from these results are as follow; (a) the sample of the study is not representative, maybe the picture would be different if a large-scale sample from across Namibia could be embarked on, (b) in Namibia, students from poor backgrounds seem to do better in school than their peers from wealthier backgrounds, (c) the same cannot be said for the students from wealthy backgrounds, as they appear to lack the motivation to do well in school.

To understand this phenomenon further, we divided the students' SES into two distinct variables. Students' SES backgrounds were recoded into low and high SES level based on their parents' educational level. Low SES included parents who did not obtained bachelor degrees and higher, while the high SES refers to parents who obtained a bachelor degree and higher. Independence sample t-test showed that there was no significant difference between low and high

SES students' performance in IR. However, low SES students were found to have performed significantly better than the high SES students in SR, ($M_{diff}=4.76$; $t=2.38$, $p<.05$).

5.4.4 Discussions

The aims of this study were to find the relationships between reading comprehension skills, inductive reasoning and scientific reasoning skills. The relationship of how RC and SES can influence students' achievement in IR and LCTSR was also explored. It was conducted in a context where SES has been found to strongly impact students' school achievement and due to societal and historic reasons English played a uniquely different role.

This study revealed that students' performance in all the three tests was above average, given the fact that these cognitive and psychological tests were new in Namibia. Therefore, we can conclude that the instruments were reliable and valid and can be used to in further research in Namibia. The results indicate that, students achieved significantly better in reading comprehension tests than in inductive and scientific reasoning tests. This finding is in line with one of the international assessments conducted in Hungary, 2015, where students' achievements reflected more equal opportunities for young learners to study in the more desirable English language associated with high cultural capital. However, another surprising result in this study was that students from low social economic background outperformed their peers from the high social economic background in LCTSR, which does not conform to the literature from the developed world. More research is needed where a representative sample is used in order to determine whether low SES students are better than high SES students in Namibia.

Small-scale studies like the present one is useful for drawing the larger picture, for testing models, and offering insights into how things operate in Namibia. As alluded to earlier, no data in the literature exists about the structure of inductive reasoning, scientific reasoning, reading comprehension skills and scientific inquiry of the population of Namibian students. Further research is needed to explain the relationships found in the present study.

As it was pointed out, inductive reasoning seems to play a more decisive role in reading comprehension achievements in English, but not in scientific reasoning. Whether this is a result of classroom methodology or is related to the linguistic features of the target languages should be the focus of further research on learnability of the second language. Furthermore, in this present study, only inductive reasoning and scientific reasoning were measured. Similar research has suggested

that, it would be useful to examine how deductive reasoning, phonemic awareness and memory interact in the development of listening comprehension, speaking and writing skills (Nikolov & Csapó, 2018). There are so many under-researched areas in the Namibian context, but these areas are also under-researched internationally and should offer new insights into how the variables studied interact.

CHAPTER 6 CONCLUSIONS, RECOMMENDATIONS, IMPLICATIONS, AND LIMITATIONS

6.1 Introduction

Through the use of four complementary sub studies this research provided insights into some of the 21st century issues and trends needed by students. Sub study 1 through 3, went further and investigated the development of the reasoning and inquiry skills. Since Namibia does not take part in the international assessment programs, (e.g. PISA, TIMSS), the aim was to ascertain students' abilities in scientific reasoning, scientific inquiry and inductive reasoning skills. As the increasing gap between girls and boys is a concern in international assessments, this study included an analysis of gender differences. Another focus of special interest was whether students' language skills (medium of instruction) influenced their achievement in SR, SI and IR.

6.2 Sub study 1 online assessment of the 5th and 7th graders' SR, IR and motivation to learn science

I begin by summarising the findings related to the general research aims from each sub studies. Study one, was the first attempt at initiating online assessment in Namibia at primary school level. Although the literature suggests that technology-based assessment may provide schools and teachers with a user-friendly instrument for monitoring the development of students' thinking skills (Pásztor et al., 2015), the same is not true for Namibia, as many public schools do not have a functional ICT infrastructure for the online assessments to be an everyday school practice.

Furthermore, the psychometric analyses of the reasoning skills tests revealed that the SR test needed to be revised in order to carry out more reliable assessments in this age group. The IR test reliability indices were high, but this was undermined by the students' poor performance. Therefore, hypotheses (H₁-H₃) were confirmed. The Rasch analyses revealed that the tests were beyond the capabilities of the primary school students. The task analysis revealed that students were not able to solve simple tasks nor find the relevant rules in the IR test. This result highlights the need to train and expose students explicitly to a different type of reasoning and thinking skill. The results revealed the fact that both SR and IR develop with age, as students from higher grade samples performed significantly better than their younger counterparts. The non-significant

correlations found between self-efficacy and test performance, might be attributed to language proficiency, here our hypotheses (H_6) was not confirmed. Self-efficacy is most likely a new word to most primary school students. The weak correlations between thinking skills in science and motivation and the relatively high average scores on motivation suggest that regardless of the students' level in reasoning skills this did not deter them from wanting to study science. The results of this study made me realize that I should move to the secondary phase with sub study 2.

6.3 Sub study 2 online assessment of scientific inquiry skills of the 9th and 11th grade students

The aim of sub study 2, was to pilot the on-line instrument for the assessment of scientific inquiry skills using the cognitive skills (scientific reasoning skills) of 9th and 11th grade students. The online assessment instrument for scientific inquiry skills proved to be reliable for the whole test hence our hypothesis (H_7) was confirmed. However, at the subscales level, three subscales that yielded (Cronbach alpha of $< .65$), indicated that it was necessary to improve the subscales in order to measure the same constructs in a more reliable way. Nonetheless, the first step was to make the instrument suitable for everyday school practice and for possible large-scale assessments in Namibia.

Strong positive correlations were found among the subscales, one of the hypothesis (H_{11}) has been confirmed. This means that with further development, the test could be used to effectively assess the abilities of students' scientific inquiry skills. Furthermore, students' mean performance was moderate but it could be improved with further training in scientific inquiry skills during their teaching and learning. One needs to ask whether the teaching and learning of science in Namibia is preparing the students thoroughly for the 21st century. Or does it concentrate on syllabi and content where the enhancement and inculcating of inquiry skills is minimal? No significant differences in performance between the two grades and genders were found, these results did not confirm the H_8 , but confirmed the H_9 . With regard to gender differences, research has indicated that gender does not influence students' understanding of nor their attitude toward science. These findings resonate with other research such as research by (Piraksa et al., 2013) in India who also found that gender does not significantly influence students' scientific reasoning ability. Therefore, the findings point to the fact that the improvement of students' scientific inquiry skills and reasoning ability is critical. This also implies that instructional pedagogy in science classroom

should place more emphasis on how to: (i) reason casually based on hypothesis generation, and (ii) design well developed science experiments, in order to enhance the development of students' scientific inquiry skills and reasoning ability.

Literature has also revealed that parents' education predicts children's educational outcomes, alongside other family characteristics such as family income, parents' occupations, and residence location. Interestingly, in this study, the results did not confirm our hypothesis, (H_{10}) as it indicate that the opposite is true in the sample assessed in Namibia. Students whose mothers did not finish primary education performed significantly better than those students whose mothers who have secondary education and beyond. My personal opinions are that being under-privileged in the Namibian context encourages students to try harder to get out of the poverty cycle, thereby outperforming their wealthier peers who do not have to strive as hard for a change in social status.

6.4 Sub study 3 the relationships and assessment of SR, SI and IR with students' SES

The purpose of study 3 was to assess the 10th and 12th grade students' abilities in scientific reasoning, inductive reasoning and scientific inquiry skills in Namibia, as well as to examine how the background variables (gender, grade, parents' level of education) affect their performances in these three domains. The paper and pencil assessment method proved to be very reliable as all three tests yielded very high internal consistency of Cronbach alpha > .87 in both age groups, therefore, our hypotheses (H_{12} & H_{13}) were confirmed. With regard to scientific reasoning, the original Lawson Classroom Test of Scientific reasoning (LCTSR) was used (Lawson, 1978, 2000). The reliability results matched with what was found by Lawson (Cronbach alpha = .78; .77 & .80) respectively among the same age group in the USA. However, the findings suggest that hypothesis testing is a more advanced ability that students begin to develop in high school. It is also the most rapidly changing (discriminating) ability amongst all the skill dimensions of LCTSR. On this skill dimension, the ceiling effect of the Lawson's test questions is a little over 80%. Two potential causes of this low ceiling have been observed in research (Han, 2013). One is the length of reading time required, which often causes students to lose track of the relevant experimental structure and variables and thereby misinterpret the questions. The other is related to the contextual elements of the questions. Students often tried to use their prior knowledge about red cells in question pair of 21-22 to answer the questions rather than using reasoning. In question pair 23-24, where a plastic bag that is semi-permeable is used goes against common sense as most plastic bags encountered

in real life are waterproof. Therefore, some students may think that the designs may be implausible, which stopped them from any further reasoning in the question.

Consequently, the SI test also proved to be highly reliable. The results matched with what the proponents of scientific inquiry skills found (Wenning, 2007). As in LCTSR, among the seven subscales of scientific inquiry skills tests, two (hypothesis identification and RQs) subscales yielded a low reliability of Cronbach alpha $< .70$, which is the acceptable Cronbach alpha in the testing field. This suggests that the items in the two subscales may not necessarily measure the same construct. Therefore, a rearrangement of items is necessary.

Mean performance results indicated that no significant difference was found between genders in all three tests, and this confirmed hypothesis 14. Concerning age groups, no significant differences were found in performance of scientific reasoning. However, the older age group performed significantly better than the younger age group in scientific inquiry skills and in inductive reasoning test. The Rasch analysis of students' ability level showed that all the three tests proved to be a good match between item difficulty and the students' abilities.

SEM analysis revealed the relationships between LCTSR, IR & SI. The SEM model for all three tests showed a good model fit, at the whole samples, and at the two-different age group. As in the literature, a positive correlation between IR and SI was found (Korom et al., 2017). However, LCTSR did not correlate significantly with IR and SI. This is a bit surprising as in the literature (Csapó, 1997; Molnár et al., 2015; Korom et al., 2017; Pásztor et al., 2017) indicated that domain general thinking skills correlate and predict most of the domain specific reasoning skills, therefore, our hypothesis (H_{16}) was not confirmed. This presents an opportunity for further instructions in this area. In sub study 1, where the online assessment method was used, strong positive correlation between IR and SR was found (Kambeyo & Csapó, 2017).

Again, in study 3, parents' educational level predicted students' achievement in a contradictory way. Students from low SES outperformed those students from high SES, and this confirmed our hypothesis (H_{15}) as per the previous studies. This creates a good premise for further research in order to find out why this is so and whether it is true for all Namibian students using a longitudinal study with a representative sample.

6.5 Sub study 4 the relationships between RC, SR, IR and SES

During the analysis of the previous studies, a possibility arose that perhaps students' performance was affected by the language component of the study, since English is not their mother tongue. This assumption led to a supplementary study being carried out. The aims of sub study four, were to investigate the relationship between reading comprehension skills, inductive reasoning and scientific reasoning skills. The effect of how RC and SES influences students' achievement in IR and SR was also explored. It was conducted in a context where SES have been found to have a big impact on students' school achievement and where English played a uniquely different role.

This study revealed that students performed above average in all three tests, especially given the fact that these cognitive and psychological tests were the first of their kind in Namibia, and this confirmed (H₁₇). The results indicate that students achieved significantly better in the reading comprehension test than in inductive and scientific reasoning tests, which implies that students have no difficulty in following instructions in English. This finding is in line with one of the international assessments conducted in Hungary, 2015, where students' achievements reflected a good grasp of English. Additionally, strong positive correlations were also found between the reading comprehension test and the two reasoning skills tests (IR & SR) thereby confirming our hypothesis 19. Consistent with the results of the two previous sub studies, the hypothesis about students from low SES performing better than students from high SES was also confirmed in this study whereby students from low social economic background significantly outperformed their peers from the high social economic background in SR, which is not line with literature from the developed world. More research is needed where a representative sample is used in order to determine whether low SES students would perform better than high SES students in Namibia.

6.6 Recommendations

In light of the findings, the following recommendations are proposed:

The study recommends that a nation-wide study be conducted on how many public schools have access to ICT. As the research suggests that technology-based assessment may provide schools and teachers with a user-friendly instrument for the monitoring the development of students' thinking and reasoning skills it is essential that the schools are equipped to implement this assessment. Explicit training on SR, SI and IR is recommended for both teachers and students

in order to improve the reasoning abilities and by extension the performance in the STEM subjects. Results from study one and two suggest that some questions might have been left unanswered like the relationship between the content of the tests and the Namibian 4-7 science curriculum. Further research is suggested in order to reveal the relationship between the content of the tests and the curriculum. Another suggestion is that a longitudinal study to find out whether the fact that students from low SES performed significantly better than students from high SES holds true for all Namibian. It would also be useful to examine how deductive reasoning, phonemic awareness and memory interact in the development of listening comprehension, speaking and writing skills. There are many under-researched areas in the Namibian context, but these areas are also under-researched internationally and should offer new insights into how the variables studied interact.

Results of the present analyses as well as those of earlier studies clearly indicate the complex interactions between students' language choice, inductive reasoning, SES, and reading comprehension achievements in L1 and L2. Small-scale studies like the present one is useful for sketching the bigger picture, for testing models, and for offering insights into the state of assessment in Namibia. As alluded to earlier, no data exists in the literature about the structure of inductive reasoning, scientific reasoning, reading comprehension skills and scientific inquiry of the population of the Namibian student. Further research is needed to explain the relationships found in the present study. As was pointed out, inductive reasoning seems to play a more decisive role in reading comprehension achievement in English, but not in scientific reasoning. Whether this is a result of classroom methodology or is as a result of data collection procedures, or may be is related to the linguistic features of the target languages could be the focus of further research on the learnability of the second language.

The study raises some questions that need to be answered, for example the socio-economic status of students and its impact on achievement in Namibia. A more representative research is encouraged to further explore this phenomenon. Further research could reveal how different item types (open- or closed-ended tasks) affect these relations or what the contribution of domain general reasoning abilities is in explaining mathematical achievement. Additionally, it would also be worth examining relationships with other domains, such as science or English reading. Students' motivation, and how it is mediated by parents and teachers, other components of their aptitude, L1 and L2 skills should also be studied, especially in the area of speaking. Last but not

the least, a longitudinal study with a more representative sample is recommended in order to ascertain the reasoning and thinking abilities in the three constructs studied here.

6.7 Educational implications

This study was one of the first attempts to carry out the assessment of SR, SI, and IR in Namibia from primary school to secondary school level. The results indicate that assessment may provide schools and teachers with a useful user-friendly instrument for observing the development of students' thinking and reasoning skills. Another important implication of this work is the use of scientific reasoning as an assessment tool. By using scientific reasoning for formative assessment, teachers can prompt students to reflect on and assess their theories (prior knowledge) as well as to expand on the reasons underlying their theories. Besides its application in the theoretical aspect of science, scientific reasoning is also applicable to experimental work. Laboratory work may be structured to be open-ended, students face a situation where theory and experimental evidence are in conflict, and they need to conduct experiments to resolve the issue. Scientific reasoning and scientific inquiry come into play when students must reconcile their theory in light of the experimental results (i.e. the evidence). The items tested in this study could be used to assess students' abilities in scientific thinking and reasoning using inquiry methods.

According to this diagnosis of students' abilities, specific learning environments can be created to promote students' inquiry skills. One goal of science education is to produce scientifically literate people with both content knowledge and knowledge of inquiry methods (Lederman & Lederman 2012), which is achievable. In school, content knowledge structured according to specific subject topics is not the only way of teaching science subjects, as this subjects-specific content knowledge could also be structured according to inquiry methods. Because no method seems to be more complex for students at this age, teachers can choose the method that fits the topic in order to foster the students' abilities (Nowak et al., 2013).

Consequently, the results suggest more representative studies should be carried out nationally in order to gauge the level of the Namibian students in SR, SI and IR. The reason for this is that this sample of the study was drawn only from the northern regions of Namibia.

Based on the two online studies and on the general current research community, research has highlighted the advantages of technology-based assessment, such as online test administration and automatic calculation of scoring which reduced the time and cost of the testing process

(Pásztor et al., 2015). Considering these advantages, the first step to be undertaken should be to make the instrument suitable for everyday school practice and for possible large-scale assessments in Namibia. By doing this, teachers would be availed with the opportunity to track their students' progress in developing the skills they require to function effectively in society today.

The question yet to be answered is, are the stakeholders in the education fraternity in Namibia ready and able to improve the ICT infrastructures in most public schools as per their blueprint? (MoE, 2001-2006). My experience shows that most schools are a long way from being ICT literate or functional (Kambeyo, 2017; Kambeyo & Csapó, 2017). Based on the results of study one and two, the possibilities for developing an online evaluation system should be further explored in the near future.

6.8 Limitations of the studies

With regard to study one and two, online assessment and computer usage might be new to most of the students in Namibia. Since they may not be familiar with technology-based assessment, this research instrument may not be able to give accurate ability and developmental level of the students. Another limitation is related to sample size in the current study. The sample is relatively small and does not represent average students within the country. Overall, in order to get an accurate picture of Namibian students' abilities in SR, SI and IR and provide better explanations, a larger sample from different regions across Namibia is advised. Such an approach can help determine students' ability and developmental level in the three constructs that have been studied and the differences between regions could be determined.

6.9 Summary

In this chapter, the conclusions, recommendations, studies implications and limitations were presented. The main aim of the study was to ascertain the students' ability level in SR, SI and IR skills and how parents' level of education (SES) influenced students' performance in the three constructs assessed. The studies revealed that the 5th and 7th grade students required more training before being able to be assessed in SR and IR. Despite this, students reported that they are motivated to continue studying science. Sub study 2 revealed that the SI instrument is reliable, and students performed slightly above average. Sub study 3 and 4 indicate that students' performance is not affected by language proficiency. Strong correlations were found between IR, RC and SI

and a weak correlation between IR and SR. Further to that students from low SES outperformed their peers from high SES. This is in direct contrast to findings in developed countries. The results of the study suggest a longitudinal approach with representative samples in order to fully ascertain and understand the Namibian students' ability in SR, SI and IR should be undertaken. It would also be interesting to confirm if students' SES influences their performance the same way as this current study revealed.

PUBLICATIONS RELATED TO THE DISSERTATION

- Julius, H. L. & Kambeyo, L. (2016). Investigating the motivating factors that influenced the University of Namibia first year undergraduate student teachers to choose teaching as a career: A case study. In H. Miranda (Ed.), *Proceedings of the 5th UNAM Annual Educational Conference* (pp. 167-168, 22-24 August 2016, Ongwediva, Namibia).
- Kambeyo, L., & Wu, H. (2018). Online assessment of students' inductive reasoning skills abilities in Oshana Region, Namibia. *International Journal of Educational Sciences*, 20(1), 137-147.
- Kambeyo, L., & Wu, H. (2017). Nemzetközi tapasztalatok a tanulói kompetenciák mérésére. In G. Hunyady, B. Csapó, G. Pusztay & J. Szivák (szerk.), *Az oktatás korproblémái.* (pp 179-189) ELTE Eötvös Kiadó, Budapest.
- Kambeyo, L., & Csapó, B. (2017). The feasibility of online assessment of students' inductive reasoning skills abilities in Namibia. *Proceedings of the 9th EDULEARN17 Conference* 3rd-5th July 2017, Barcelona, Spain.
- Kambeyo, L., & Csapó, B. (2018). Scientific reasoning skills: A theoretical background on science education. *Reform Forum Journal for Educational Reform in Namibia*, 26(1), 29-39.
- Kambeyo, L. (2016). Scientific reasoning skills: A theoretical background on science education. In H. Miranda (Ed.), *Proceedings of the 5th UNAM Annual Educational Conference* (pp. 192-196), 22-24 August 2016, Ongwediva, Namibia.
- Kambeyo, L. (2017) Assessing students' abilities of scientific inquiry skills. 6th *Interdiszciplináris Doktorandusz Konferencia (IDK2017)* (6th Interdisciplinary Doctoral Conference), Pecs, Hungary, 19-21 May, 2017.
- Kambeyo, L. (2017). Online assessment of scientific reasoning skills: The case of Namibia. 15th *Conference on Educational Assessment (CEA)*, Szeged, Hungary, 6-8 April, 2017.
- Kambeyo, L. (2017). The possibilities of assessing students' Scientific Inquiry Skills Abilities Using an Online Instrument: A Small-Scale Study in the Omusati Region, Namibia. *European Journal of Educational Sciences*, 4(2), 1-21.
- Kambeyo, L. (2018). The assessment of English reading comprehension and its predictive power on inductive reasoning. 16th *Conference on Educational Assessment (CEA)*, 26-28 April 2018, Szeged, Hungary.
- Kambeyo, L. (2018). The assessment of scientific reasoning, inductive reasoning and scientific inquiry using paper and pencil in Namibia. *Proceedings of 10th EDULEARN18 Conference*, 2nd-4th, July 2018, Palma de Mallorca Spain. Paper Accepted
- Kambeyo, L. (2018). The assessment of scientific reasoning, inductive reasoning and scientific inquiry using paper and pencil in Namibia. 16th *Conference on Educational Assessment (CEA)*, 26-28 April, 2018, Szeged, Hungary.

- Kambeyo, L., Pásztor, A., Korom, E., & Németh, B. M. (2018). Online assessment of thinking skills in science context and motivation to learn science among the grade 7 students in Oshana region, Namibia. *26th Annual Conference of the Southern African Association for Research in Mathematics, Science and Technology Education (SAARMSTE)* (p. 129), Gaborone, Botswana, 16-19 January 2018.
- Kambeyo, L., Pásztor, A., Korom, E., Németh, B. M., & Csapó, B. (2017). Online assessment of scientific reasoning and motivation to learn science: A pilot study in Namibia. *17th Biennial Conference of the European Association for Research on Learning and Instruction (EARLI)* (p. 129), Tampere, Finland, August 29- September 2, 2017.
- Orosz, Z. G., B. Németh, M., Kambeyo, L., Korom, E. (2017). Assessing inquiry skills in Hungary and Namibia. In J. Kerülő, T. Jenei, I. Gyarmati (Eds.), *ONK 2017. XVII. National Educational Conference* (p. 532). Nyíregyháza, Hungary, 9-11 November, 2017.
- Wu, H., Kambeyo, L., & Molnár, G. (2017). The feasibility and reliability of computer-based assessment of inductive reasoning in Namibia and China: A comparison study. *15th Conference on Educational Assessment (CEA)*, 6-7 April 2017, Szeged, Hungary.

REFERENCES

- Adams, R. J., & Wu, M. L. (2002). *PISA 2000 technical report*. Paris: OECD.
- Adey, P. S., & Shayer, M. (1994). *Really raising standards*. London: Routledge.
- Adey, P., & Csapó, B. (2012). Developing and assessing scientific reasoning. In B. Csapó & G. Szabo, *Framework for diagnostic assessment of science* (pp. 17-49). Budapest: Nemzeti Tankönyvkiado.
- Allard, D. & Barman, C. (1994). The learning cycle: A sound alternative to current college science teaching. *BioScience*, 44(2), 99-101
- Allchin, D. (2011). Evaluating knowledge of the nature of (whole) science. *Science Education*, 95, 518-542.
- Almond, R. G., & Mislevy, R. J. (1999). Graphical models and computerized adaptive testing. *Applied Psychological Measurement*, 23, 223-237.
- Al-Zoubi, T., El-shar'a, I., & Al-Salam, M. K. (2009). The scientific reasoning level of students' in the Faculty of Science in Al-Hussein Bin Talal University and its affection of gender, teaching level, and specialization. *An-Najah University Journal for Research - Humanities*, 23(2), 401-437.
- American Association for the Advancement of Science. (1990). *Science for all Americans*. New York: Oxford University Press.
- American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. New York: Oxford University Press.
- American Association of Colleges and Universities. (2007). *College learning for the new global century*. Washington, DC: AACU.
- Anderson, R. D. (2007). Inquiry as an organizing theme for science curricula. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 807-830). Mahwah, NJ: Lawrence Erlbaum Associates.
- Asad khan, R. M., Iqbal, M., & Tasneem, S. (2015). The influence of parents' educational level on secondary school students' academic achievements in district Rajanpur. *Journal of Education and Practice*, 6(16), 76-79.
- Baker, C. (1996). *Foundations of bilingual education and bilingualism*. Clevedon, Avon: Multilingual Matters.
- Baker, E. L., & Mayer, R. E. (1999). Computer-based assessment of problem solving. *Computers in Human Behavior*, 15, 269-282.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, 52, 1-26.
- Bandura, A. (2006). Going global with social cognitive theory: From prospect to pay dirt. In S. I. Donaldson, D. E. Berger & K. Pezdek (Eds.), *The rise of applied psychology: New frontiers and rewarding careers* (pp. 53-70). Mahwah, NJ: Erlbaum.

- Bandura, A., & Locke, E. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology*, 88, 87-99.
- Bao, L., & Koenig, K. (2012). *TI21: A technology enhanced inquiry framework for developing and assessing 21st century skills*. Available from iSTARAssessment.org
- Bao, L., Cai, T., Koenig, K., & Fang, K. (2009). *Learning and scientific reasoning*. Washington, DC: AAAS.
- Beaumont-Walters, Y., & Soyibo, K. (2001). An analysis of high school students' performance on five integrated science process skills. *Research in Science & Technological Education*, 19, 133-145.
- Bennett, R. E. (1999). Using new technology to improve assessment. *Educational Measurement: Issues and Practice*, 18, 5-12.
- Bennett, R. E., Persky, H., Weiss, A. R., & Jenkins, F. (2007). *Problem solving in technology- rich environments*. A report from the NAEP Technology-Based Assessment Project, Research and Development Series. NCES 2007-466.
- Bialystok, E. (2002). Acquisition of literacy in bilingual children: A framework for research. *Language Learning*, 52, 159-199.
- Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., & Rumble, M. (2012). Defining 21st century skills. In P. Griffin, B. McGaw & E. Care (Eds.), *Assessment and teaching of the 21st century skills* (pp. 17-66). New York: Springer. Retrieved from <http://atc21s.org/wp-content/uploads/2017/03/25-Defining-21st-Century-Skills.pdf>
- Bisanz, J., Bisanz, G., & Korpan, C. A. (1994). Inductive reasoning. In R. Sternberg (Ed.), *Thinking and problem solving* (pp. 181-213). San Diego: Academic Press.
- Black, A. E., & Deci, E. L. (2000). The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: A self-determination theory perspective. *Science Education*, 84, 740-756.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Boyer, T. W., Levine, S. C., & Huttenlocher, J. (2008). Development of proportional reasoning: Where young children go wrong. *Developmental Psychology*, 44(5), 1478-1490.
- Brotherton, P. N., & Preece, P. F. W. (1995). Science process skills: Their nature and interrelationships. *Research in Science & Technological Education*, 13, 5-11.
- Bryn, R. R., Glyn, S. M., & Kittleson, J. M. (2011). Motivation, achievement and advanced placement intent of high school students learning science. *Journal of Science Education*, 1049-1065. Published online.
- Bunce, D.M., & Hutchinson, K.D., (1993). The use of the GALT (Group Assessment of Logical Thinking) as a predictor of academic success in college chemistry. *Journal of Chemical Education*, 70, 183-187.
- Bunderson, C. V., Inouye, D. K., & Olsen, J. B. (1989). The four generations of computerized educational measurement. In R. L. Linn (Ed.), *Educational measurement* (pp. 367-407). Washington, DC: American Council on Education.

- Bybee, R. W. (2000). Teaching science as inquiry. In J. Minstrell & E. H. van Zee (Eds.), *Inquiring into inquiry learning and teaching in science* (pp. 20-46). Washington, DC: American Association for the Advancement of Science.
- Bybee, R. W. (2009). *The BSCS 5E instructional model and 21st century skills*. National Academies Board on Science Education, Washington, DC: Retrieved March 25, 2017, from: https://www7.nationalacademies.org/bose/1Bybee_21st%20Century_Paper.pdf
- Čandrlić, S., Katić, M. A., & Dlab, M. H. (2014). Online vs. paper-based testing: A comparison of test results. *37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)* (pp. 775-778), Opatija, Croatia, 26-30 May 2014.
- Capps, D. K., & Crawford, B. A. (2013). Inquiry-based instruction and teaching about nature of science: Are they happening? *Journal of Science Teacher Education*, 24, 497-526.
- Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education*, 86, 175-218.
- Clariana, R. B., & Wallace, P. E. (2002). Paper-based versus computer-based assessment: Key factors associated with the test mode effect. *British Journal of Educational Technology*, 33 (5), 595-604.
- Clark, R. E. (1994). Media will never influence learning. *Educational Technology, Research and Development*, 42(2), 21-29.
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (7th ed.). London: Routledge.
- Coil, C., Wenderoth, M. P., Cunningham, M., & Dirks, C. (2010). Teaching the process of science: Faculty perceptions and an effective methodology. *CBE Life Sciences Education*, 9, 524-535.
- Coletta, V. P., & Phillips, J. A. (2005). Interpreting FCI scores, normalized gain, reinstruction scores, scientific reasoning ability. *American Journal of Physics*, 73(12), 1172-1179.
- Cook, J. L., & Cook, G. (2005). *Child development principles and perspectives*. Boston: Pearson.
- Council of Europe. (2001). *Common European framework of reference for languages: learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Csapó, B. (1990). Integrating the development of the operational abilities of thinking and the transmission of knowledge. In H. Mandl, E. De Corte, N. Bennett, & H. F. Friedrich (Eds.), *Learning and instruction. European research in an international context. Volume 2.2. Analysis of complex skills and complex knowledge domains* (pp. 85-94). Oxford, Pergamon Press.
- Csapó, B. (1997). Development of inductive reasoning: Cross-sectional measurements in an educational context. *International Journal Behavioral Development*, 20(4), 609-625.
- Csapó, B. (2010). Research on development in educational contexts: The Hungarian Educational Longitudinal Program. Keynote address. *Children and Youth in Changing Societies. 3rd International Conference*. Thessaloniki, Greece, December 2-4, 2010. 170.
- Csapó, B. (2012). Developing a framework for diagnostic assessment of early science. In S. Bernholt, K. Neuman, & P. Nentwig, *Making it tangible: Learning outcomes in science education* (pp. 55-78). Berlin: Waxmann.

- Csapó, B., & Funke, J. (2017). The development and assessment of problem solving in 21st-century schools. In B. Csapó & J. Funke (eds.), *The Nature of Problem Solving: Using Research to Inspire 21st Century Learning*, OECD Publishing, Paris. DOI: <http://dx.doi.org/10.1787/9789264273955-3-en>
- Csapó, B., & Nikolov, M. (2009). The cognitive contribution to the development of proficiency in a foreign language. *Learning and Individual Differences*, 19, 203-218. doi:10.1016/j.lindif.2009.01.002
- Csapó, B., & Szabo, G. (2012). *Framework for diagnostic assessment of science*. Budapest: Nemezeti Tankönyvkiadó.
- Csapó, B., Ainley, J., Bennett, R., Latour, T., & Law, N. (2012). Technological issues of computer-based assessment of 21st-century skills. In B. McGaw, P. Griffin & E. Care (Eds.), *Assessment and teaching of 21st-century skills* (pp. 143-230). New York: Springer.
- Csapó, B., Lorincz, A., & Molnár, G. (2012). Innovative Assessment Technologies in Educational Games Designed for Young Students. In D. Ifenthaler, D. Eseryel, & X. Ge (Eds.), *Assessment in game-based learning: foundations, innovations, and perspectives* (pp. 235-254). New York: Springer. <http://dx.doi.org/10.1007/978-1-4614-3546-4>
- Csapó, B., Molnár, G., & Nagy, J. (2014). Computer-based assessment of school readiness and early reasoning. *Journal of Educational Psychology*, 106(3), 639.
- Csapó, B., Molnár, G., & Tóth, K. (2009). Comparing paper-and-pencil and online assessment of reasoning skills. A pilot study for introducing electronic testing in large-scale assessment in Hungary. In F. Scheuermann & J. és Björnsson, *The transition to computer-based assessment: New approaches to skills assessment and implications for large-scale testing* (pp. 120-125). Luxemburg: Office for Official Publications of the European Communities.
- Cummins, J. (2000). *Language, power and pedagogy: Bilingual children in the crossfire*. Clevedon, Avon: Multilingual Matters.
- Darling-Hammond, L. (1994). Performance-based assessment and educational equity. *Harvard Educational Review*, 64(1), 5-29.
- De Boer, G. E. (2004). Historical perspectives on inquiry teaching in schools. In L. B. Flick & N. G. Lederman (Eds.), *Scientific inquiry and nature of science* (pp. 17-35). Dordrecht: Kluwer Academic Publishers.
- De Jager, T., & Ferreira, J. G. (2003). Factors preventing the development of process skills of biology secondary school learners in South Africa. *Educare*, 32, 186-198.
- De Konig, E. (2000). *Inductive reasoning in primary education: Measurement, teaching, transfer*. Zeist: Kerckebosch.
- De Koning, E., & Hamers, J. H. M. (1999). Teaching inductive reasoning: Theoretical background and educational implications. In J. H. M. Hamers, J. E. H. van Luit, & B. Csapó (Eds.), *Teaching and learning thinking skills* (pp. 156-188). Lisse: Swets & Zeitlinger.
- De Koning, E., Hamers, J. H. M., Sijtsma, K., & Vermeer, A. (2002). Teaching inductive reasoning in primary school. *Developmental Review*, 22, 211-241. Doi:10.1006/drev.2002.0548

- Dede, C. (2005). Planning for “neomillennial” learning styles: Implications for investments in technology and faculty. In J. Oblinger and D. Oblinger (Eds.), *Educating the net generation*, (pp. 226-247). Boulder, CO: EDUCAUSE Publishers.
- Dede, C. (2009). Immersive interfaces for engagement and learning. *Science*, 323(5910), 66-69.
- Dede, C. (2009). *Comparing frameworks for “21st century skills.”* Harvard Graduate School of Education. Retrieved December 19, 2016 from http://www.watertown.k12.ma.us/dept/ed_tech/research/pdf/ChrisDede.pdf
- Dimitrov, D. M. (1999). Gender differences in science achievement: Differential effect of ability, response format, and standards of learning outcomes. *School Science and Mathematics*, 99(8), 445-450.
- Driver, R., Leach, J., Millar, R., & Scott, P. (1996). *Young people’s images of science*. Bristol: Open University Press.
- Druger, M. (2006). Experiential learning in a large introductory biology course. In J. J. Mintzes & W. H. Leonard (Eds.), *Handbook of college science teaching* (pp. 37-43). Arlington, VA: National Science Teachers Association Press.
- Duggan, S., Johnson, P., & Gott, R., (1996). A critical point in investigative work: Defining variables. *Journal of Research in Science Teaching* 33 (5): 461-474.
- Eccles, J. S., & Davis-Kean, P. E. (2005). Influence of parents’ education on their children’s educational attainments: The role of parent and child perceptions. *London Review of Education*, 3(3), 191-204.
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53, 109-132.
- Eccles, J. S., Simpkins, S. D., & Davis-Kean, P. E. (2006). Math and science motivation: A longitudinal examination of the links between choices and beliefs. *Developmental Psychology*, 42, 70-83.
- Educational Testing Services (2007). *Education issues: Listening, Learning, & Leading*. Washington DC.
- Erduran, S. & Msimanga, A. (2014). Science curriculum reform in South Africa: Lessons for professional development from research on argumentation in science education. *Education as Change*, 18, S33-S46.
- Federico, P. A. (1989). Computer-based and paper-based measurement of recognition performance. *Navy Personnel Research and Development Center Report NPRDC-TR-89-7*. Available from ERIC: ED 306 308.
- Feinstein, N. (2011). Salvaging science literacy. *Science Education*, 95, 168 -185.
- Fenichel, M., & Schweingruber, H. (2010). *Surrounded by science: Learning science in informal environments*. Washington, DC.: National Press.
- Fowler, J. W. (1981). *Stages of faith: The psychology of human development and the quest for meaning*. New York: Harper & Row.
- Fu, A. C., Shavelson, R. J., & Raizen, S. A. (2009). *The nation report card: A vision of large-scale science assessment*. Washington, DC: AAAS.

- Furtak, E. M., Hardy, I., Beinbrech, T., Shavelson, R. J., & Shemwell, J. T. (2008). *A framework for analyzing reasoning in science classroom discourse*. New York: AAAS.
- Gaigher, E., Lederman, N., & Lederman, J. S. (2014). Knowledge about inquiry: A study in South African high schools, *International Journal of Science Education*, 36(18), 3125-3147. DOI: 10.1080/09500693.2014.954156
- Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 199-241). London: Cambridge University Press.
- Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15, 1-38.
- Gilbert, J. K., Pietrocola, M., Zylbersztajn, A., & Franco, C. (2000). Science and education, Notions of reality, theory and models. In Gilbert, J. K and Boulter, C. J. (Eds.), *Developing models in science education* (pp. 19-40). Dordrecht: Kluwer.
- Gillies, R. M., Nichols, K., Burg, G., & Haynes, M. (2014). Primary students' scientific reasoning and discourse during cooperative inquiry-based science activities. *International Journal of Education Research*, 63, 127-140.
- Glynn, S. M., & Koballa, T. R. Jr. (2006). Motivation to learn college science. In J. J. Mintzes & W. H. Leonard (Eds.), *Handbook of college science teaching* (pp. 25-32). Arlington, VA: National Science Teachers Association Press.
- Glynn, S. M., Brickman, P., Armstrong, N., & Taasoobshirazi, G. (2011). Science motivation questionnaire II: Validation with science majors and non-science majors. *Journal of Research in Science Teaching*, 46, 127-146.
- Goldman, S. R., & Pellegrino, J. W. (1984). Deductions about induction: Analyses of developmental and individual differences. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 2, pp. 149-197). Hillsdale, NJ: Erlbaum.
- Goodrum, D., & Rennie, L. (2007). *Australian school science education national plan*. Canberra: ACT: Department of Education, Science and Training.
- Gormally, C., Brickman, P., & Lutz, M. (2012). Developing a test of scientific literacy skills (TOSLS): Undergraduate evaluation of scientific information and arguments. *Life Science Education*, 11(3), 364-377.
- Goswami, U. (1991). Analogical reasoning: What develops? A review of research and theory. *Child Development*, 62(1), 1-22.
- Goswami, U. (1998). *Cognition in children*. Hove, UK: Psychology Press.
- Goswami, U., & Brown, A. L. (1989). Melting chocolate and melting snowmen: Analogical reasoning and causal relations. *Cognition*, 35, 69-95.
- Gott, R., & Duggan, S. (1998). Understanding scientific evidence – Why it matters and how it can be taught. In M. Ratcliffe (Ed.), *ASE (The Association for Science Education) Guide to Secondary Science Education*, (pp. 92–99). Cheltenham: Stanley Thornes.
- Gott, R., & Duggan, S. (1995). *Investigative work in the science curriculum*. Buckingham: Open University Press.

- Greiff, S., Wüstenberg, S., Csapó, B., Demetriou, A., Hautamäki, J., Greasser, A. C., & Martin, R. (2014). Domain-general problem-solving skills and education in the 21st century. *Educational Research Review*, 13, 78-83.
- Hamers, J. H. M., De Koning, E., & Sijtsma, K. (1998). Inductive reasoning in the third grade: Intervention promises and constraints. *Contemporary Educational Psychology*, 23, 132-148.
- Hammann, M., Phan, T. T. H., Ehmer, M., & Grimm, T. (2008). Assessing pupils' skills in experimentation. *Journal of Biological Education* 42(2), 66-72.
- Han, J. B. S. (2013). *Scientific reasoning: Research, development, and assessment*. Unpublished doctoral dissertation, The Ohio State University, Columbus, USA.
- Hannaway, J., & Hamilton, L. S. (2008). *Performance based accountability policies: Implications for school and classroom practices*. Washington, DC, Washington.
- Harlen, W. (1999). Purposes and procedures for assessing science process skills: Assessment in education: *Principles, Policy and Practice*, 6, 129-144.
- Harlen, W. (2013). *Assessment and inquiry-based science education: Issues in policy and practice*. Trieste: TWAS-Strada Costiera.
- Hawkins, J., & Pea, R. D. (1987). Tools for bridging every day and scientific thinking. *Journal of Research in Science Teaching*, 24(4), 291-301.
- Hulstijn, J. (2006). *Defining and measuring the construct of second-language proficiency*. Plenary address at the Joint Conference of American Association of Applied Linguistics and Canadian Association of Applied Linguistics, Montreal, Canada, June, 2006.
- Hulstijn, J. (2015). *Language proficiency in native and non-native speakers: Theory and research*. Amsterdam: John Benjamins Publishing Company.
- Ipinge, S. M., & Likando, G. N. (2012). The educational assessment reforms in post-independence Namibia: A critical analysis. *SA-eDUC JOURNAL*, 1-10.
- Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking*. London: Routledge Kegan Paul.
- International Society for Technology in Education (2007). *Accelerate innovation, transform learning*. New York.
- Jenkins, H., Clinton, K., Purushotma, R., Robson, A. J., & Weigel, M. (2010). *Confronting the challenges of participatory culture: Media education for the 21st century*. Chicago, Illinois: MacArthur foundation Publishers.
- Justi, R. S., & Gilbert, J. K., (2002). Modelling, teachers' views on the nature of modelling, and implications for the education of modellers. *International Journal of Science Education* 24 (4): 369-387.
- Kambeyo, L., & Csapó, B. (2017). *The feasibility of online assessment of students' inductive reasoning skills abilities in Namibia*. Proceedings of 9th EDULEARN17 Conference 3rd-5th July 2017, Barcelona, Spain.
- Kambeyo, L., Pásztor, A., Korom, E., Németh, B. M., & Csapó, B. (2017). Online assessment of scientific reasoning and motivation to learn science: A pilot study in Namibia. *17th biennial conference of the European Association for Research on Learning and Instruction (EARLI)*. (p. 129), Tampere, Finland, August 29- September 2, 2017.

- Kang, S., Scharmann, L. C., Noh, T., & Koh, H. (2005). The influence of students' cognitive and motivational variables in respect of cognitive conflict and conceptual change. *International Journal of Science Education*, 27, 1037-1058.
- Keller, M. M., Neumann, K., & Fischer, H. E. (2017). The impact of physics teachers' pedagogical content knowledge and motivation on students' achievement and interest. *Journal of Research in Science teaching*, 54(5), 586-614. doi: 10.1002/tea.21378
- Kelly, G. J. (2011). Scientific Literacy, discourse, and epistemic practices. In C. Linder, L. Östman, D. A. Roberts, P. Wickman, G. Erikson, & A. McKinnon (Eds.). *Exploring the Landscape of Scientific Literacy*, (pp.61-73) Pennsylvania, Routledge.
- Klahr, D. (2000). *Exploring science: The cognition and development of discovery processes*. Cambridge, MA: MIT Press.
- Klahr, D. (2005). Early science instructions: Addressing fundamental issues. *Psychological Science*. 16(11), 871-873.
- Klauer, K. J. (1990). Paradigmatic teaching of inductive thinking. In H. Mandl, E. De Corte, S. N. Bennett & H. F. Friedrich (Eds.), *Learning and instruction. European research in an international context. Analysis of complex skills and complex knowledge domains* (pp. 23-45). Oxford: Pergamon Press.
- Klauer, K. J. (1993). *Denktraining für Jugendliche* [Training to reason for youths]. Göttingen: Hogrefe.
- Klauer, K. J. (1996). Begünstigt induktives Denken das Lösen komplexer Probleme? Experimentellen Studien zu Leutners sahel-Problem [Has inductive thinking a positive effect on complex problem solving?]. *Zeitschrift für Experimentelle Psychologie*, 43(1), 361-366.
- Klauer, K. J. (1996). Teaching inductive reasoning: Some theory and three experimental studies. *Learning and Instruction*, 6(1), 37-57.
- Klauer, K. J. (1999). Fostering higher order reasoning skills: The case of inductive reasoning. In J. H. M. Hamers, J. E. H. van Luit, & B. Csapó' (Eds.), *Teaching and learning thinking skills* (pp. 131-156). Lisse: Swets & Zeitlinger.
- Klauer, K. J., & Phye, G. (1994). *Cognitive training for children: A developmental program of inductive reasoning and problem solving*. Lisse: Swets & Zeitlinger.
- Klauer, K. J., & Phye, G. D. (2008). Inductive reasoning: A training approach. *Review of Educational Research*, 78, 85-123.
- Klauer, K. J., Wilmes, K. & Phye, G. D. (2002). Inducing inductive reasoning: Does it transfer to fluid intelligence? *Contemporary Educational Psychology*, 27(1), 1-25.
- Koballa, T. R. Jr., & Glynn, S. M. (2007). Attitudinal and motivational constructs in science education. In S. K. Abell & N. Lederman (Eds.), *Handbook for research in science education* (pp. 75-102). Mahwah, NJ: Erlbaum.
- Koda, K. (2007). Reading and language learning: Cross linguistic constraints on second language reading development. *Language Learning*, 57, Supplement 1, 1-44.

- Korom, E., Nagy, L., Nemeth, M. B., Radnoti, K., Farkas, M. A., Markocsi, I. R., Toth, Z., Csikos, Cs., & Wagner, E. (2012). Detailed frameworks for diagnostic Assessment of science. In B. Csapó & G. Szabo (Eds.): *Frameworks for diagnostic assessment of science* (pp. 175-306). Budapest: Nemezi Tankönyvkiado.
- Korom, E., Németh, M., Pásztor, A., & Csapó, B. (2017). Relationship between scientific and inductive reasoning in grades 5 and 7. *17th Biennial conference of the European Association for Research on Learning and Instruction (EARLI)* (pp. 128-129). Tampere, Finland, August 29-September 2, 2017.
- Koslowski, B. (1996). *Theory and evidence: The development of scientific reasoning*. Cambridge, MA: MIT Press.
- Kuger, S., Klieme, E., Jude, N., & Kaplan, D. (Eds.). (2016). *Assessing contexts of learning. An international perspective*. Cham, Switzerland: Springer International Publishing.
- Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, **96**, 674-689
- Kuhn, D. (2002). What is scientific thinking and how does it develop? In U. Goswami (Ed.), *Handbook of childhood cognitive development*. (pp 371-393). Malden, England: Blackwell.
- Kuhn, D. (2010). What is scientific thinking and how does it develop? In U. Goswami, *Handbook of childhood cognitive development* (pp. 497-523). Oxford: Wiley Blackwell.
- Kuhn, D., Katz, J. B., & Dean, D. (2004). Developing reason. *Thinking & Reasoning*, *10* (2), 197-219.
- Lappan, G. (2000). A vision of learning to teach for the 21st century. *School Science and Mathematics*, *100*(6), 319-325.
- Lawson, A. E. (1978). The development and validation of a classroom test of formal reasoning. *Journal of Research in Science Teaching*, *15*(1), 11-24.
- Lawson, A. E. (2000). Development and validation of the classroom test of formal Reasoning. *Journal of Research in Science Teaching*, 11-24.
- Lawson, A. E., Banks, D. L., & Logvin, M. (2007). Self-efficacy, reasoning ability, and achievement in college biology. *Journal of Research in Science Teaching*, *44*, 706-724.
- Lazonder, A. W., & Kemp, E. (2012). Bit by bit or all at once? Splitting up the inquiry task to promote children's scientific reasoning. *Learning and Instruction*, *22*(6), 458-464.
- Lederman, N. G., Antink, A., & Bartos, S. (2012). Nature of science, scientific inquiry, and socio scientific issues arising from genetics: A pathway to developing a scientifically literate citizenry. *Science and Education*. doi:10.1007/s11191-012-9503-3
- Lederman, N., & Lederman, J. (2012). Nature of scientific knowledge and scientific inquiry. In B. J. Fraser, K. Tobin, & C. J. McRobbie (Eds.), *Second international handbook of science education* (pp. 335-359). Dordrecht: Springer.
- Lee, C. Q. & She, H. C. (2010). Facilitating Students' Conceptual Change and Scientific Reasoning Involving the Unit of Combustion. *Research in Science Education*, *40* (4) 479-504.
- Lefrançois, G. R. (2001). *Of children: An introduction to child and adolescent development* (9th ed.). Belmont, CA, US: Wadsworth/Thomson Learning.

- Leggett, M., Kinnear, A., Boyce, M., & Bennett, I. (2004). Student and staff perceptions of the importance of generic skills in science. *Higher Education Research & Development*, 23, 295-312.
- Levinson, P. J., & Carpenter, R. L. (1974). An analysis of analogical reasoning in children. *Child Development*, 45, 857-861.
- Lunzer, E. A. (1965). Problems of formal reasoning in test situations. *Monographs of the Society for Research in Child Development*, 30(2), 19-46.
- Mayer, D., Sodian, B., Koerber, S., & Schwipert, K. (2014). Scientific reasoning in elementary school children: Assessment and relation with cognitive abilities. *Learning and Instruction*, 29, 43-55.
- Mayer, J. (2007). Inquiry as scientific problem solving. In D. Kruger & H. Vogt (Eds.), *Theories in biology didactic* (pp. 177-186). Heidelberg: Springer.
- Mayer, R. (2008). *Learning and instruction*. New Jersey: Pearson.
- Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.
- Mayrath, M. C., Clarke-Midura, J., Robinson, D. H., & Schraw, G. (2012). *Technology based assessments for the 21st century skills: Theoretical and practical implications for modern research*. Charlotte: Information Age Publishing.
- Mayrath, M. C., Nihalani, P. K., & Robinson, D. H. (2011). Varying tutorial modality and interface restriction to maximize transfer in a complex simulation environment. *Journal of Educational Psychology*, 10(2), 257-268. doi:10.1037/a0022369
- Mazlo, J., Dormedy, D. F., Neimoth-Anderson, J. D., Urlacher, T., Carson, G. A., & Kelter, P. B. (2002). Assessment of motivational methods in the general chemistry laboratory. *Journal of College Science Teaching*, 36, 318-321.
- Mazzeo, J., Druesne, B., Raffeld, P. C., Checketts, K. T., & Muhlstein, A. (1991). Comparability of computer and paper-and-pencil scores for two CLEP general examinations. *College Board Report No. 91-5*. Available from ERIC: ED 344 902.
- McKay, P. (2006). *Assessing young language learners*. Cambridge: Cambridge University Press.
- McShane, J. (1991). *Cognitive development: An information processing approach*. Oxford: Oxford University Press
- Mead, A. D., & Drasgow, F. (1993). Equivalence of computerized and paper-and-pencil cognitive ability tests: A meta-analysis. *Psychological Bulletin*, 114, 449-458.
- Metiri Group & NCREL. (2003). *EnGauge 21st century skills: Literacy in the digital age*. Chicago, IL: NCREL.
- Millar, R. (1989). What is 'scientific method' and can it be taught? In J. J. Wellington (Ed.), *Skills and processes in science education: A critical analysis* (pp. 47-62). London: Routledge.
- Millar, R. (1997). Science education for democracy: What can the school curriculum achieve? In R. Levinson, & J. Thomas (Eds.), *Science today: Problem or crisis?* (pp. 87-101). London: Routledge.
- Millar, R. (2002) Towards a science curriculum for public understanding. In S. Amos & R. Boohan (eds.), *Teaching Science in Secondary Schools: a Reader*, (pp. 112-128). London, Routledge.

- Millar, R. (2006). Twenty first century science: Insights from the design and implementation of a scientific literacy approach in school science. *International Journal of Science Education*, 28, 1499-1521.
- Millar, R., & Osborne, J. (1998). *Beyond 2000: Science education for the future*. London: Nuffield Foundation.
- Ministry of Basic Education, Sport and Culture. (2001). *ICT policy for education*. Windhoek: Polytechnic Press.
- Ministry of Education and Culture (1993). *Toward education for all: A development brief for education, culture and training*. Windhoek: Gamsberg Macmillan.
- Ministry of Education. (2007). *Education and training sector and Improvement program*. Windhoek: Government Press.
- Ministry of Education. (2008). *National Syllabus for Primary Education*. Windhoek: Government Press.
- Ministry of Education. (2010). *National Curriculum for Basic Education*. Okahandja: NIED.
- Ministry of Education. (2014). *Natural Science and Health education*. Okahandja: NIED.
- Ministry of Human Resources. (1996). *National Core Curriculum*. Budapest: Department of Education.
- Ministry of Human Resources. (2011). *Education in Hungary: Past, present, future-An overview*. Budapest: Department of EU Relations.
- Molnár, G. (2011). Playful fostering of 6- to 8-year-old students' inductive reasoning. *Thinking Skills and Creativity*, 6(2), 91-99.
- Molnár, G., Greiff, S., & Csapó, B. (2013). Inductive reasoning, domain specific and complex problem solving: relation and development. *Thinking skills and Creativity*, 9(8), 35-45.
- Morris, B. J., Croker, S., Masnick, A. M., & Zimmerman, C. (2015). *Intech. Open science, Open mind*. Retrieved October 14, 2015, from Creative Commons, <http://creativecommons.org>
- Muthén, L. K., & Muthén, B. O. (2010). *Mplus user's guide*. Los Angeles, CA: Muthén & Muthén.
- Nagy, J. (2009). Renewing primary education. In K. Fazekas, J. Kollo & J. Varga, *Green book: For the renewal of public education in Hungary* (pp. 61-79). Budapest: ECOSTAT.
- Nagy, L.-né., Korom, E., Pásztor, A., Veres, G., & B. Németh, M. (2015). A természettudományos gondolkodás online diagnosztikus értékelése [Online diagnostic assessment of scientific reasoning]. In B. Csapó, E. Korom, & Gy. Molnár (Eds.), *A természettudományi tudás online diagnosztikus értékelésének tartalmi keretei [Framework for the online assessment of scientific reasoning]* (pp. 35–116). Budapest: Oktatókutató és Fejlesztő Intézet.
- National Curriculum Board. (2009). *The shape of the Australian curriculum*. Canberra: Ministry of Education.
- National Institute for Educational Development. (2008). *The National Curriculum for Basic Education*. Windhoek: Ministry of Education.
- National Research Council. (1996). *National Science Education Standards*. Washington, DC: National Academies Press.

- National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts and core ideas*. Washington, DC: National Academy Press.
- National Science Teachers Association (2004) *National Science Education Standards*. Washington, DC: National Academies Press.
- Nemeth, M. B., & Korom, E. (2012). Science literacy and the application of scientific knowledge. In B. Csapó & G. Szabo, *Framework for diagnostic assessment of science* (pp. 55-87). Budapest: Nemzeti Tankönyvkiadó Zrt.
- Niaz, M., & Robinson, W.R. (1992). From 'algorithmic mode' to 'conceptual gestalt' in understanding the behavior of gases: an epistemological perspective. *Research in Science and Technological Education*, 10, 53-64.
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice, *Studies in Higher Education*, 31(2), 199-218
- NIED. (2009). *The National Curriculum for Basic Education*. Okahandja: Nied.
- NIED. (2014). *Ministry of Education. Natural Science and Health Education Syllabus*, 1-56.
- Nikolov, M., & Csapó, B. (2010). The relationship between reading skills in early English as a foreign language and Hungarian as a first language. *International Journal of Bilingualism*. 14, 315-329.
- Nikolov, M., & Csapó, B. (2018). The relationships between 8th graders' L1 and L2 readings skills, inductive reasoning and socio-economic status in early English and German as a foreign language programs. *System*, 73, 184-206.
- Nikolov, M., & Józsa, K. (2006). Relationships between language achievements in English and German and classroom-related variables. In M. Nikolov. & J. Horváth (Eds.), *UPRT 2006: Empirical studies in English applied linguistics* (pp. 197-224). Pécs: Lingua Franca Csoport. PTE.
- Nikolov, M., & Szabó, G. (2015). A study on Hungarian 6th and 8th graders' proficiency in English and German at dual-language schools. In D. Holló & K. Károly (Eds.), *Inspirations in foreign language teaching: Studies in applied linguistics. Language pedagogy and language teaching* (pp. 184-206). Harlow: Pearson Education.
- Nowak, K. H., Nehring, A., Tiemann, R., & Upmeyer zu Belzen, A. (2013) Assessing students' abilities in processes of scientific inquiry in biology using a paper-and-pencil test. *Journal of Biological Education*, 47(3), 182-188. DOI: 10.1080/00219266.2013.822747
- OECD. (2005). *The definition and selection of key competencies: Executive summary*. Paris.
- OECD. (2013). *PISA 2015 Draft Science Framework*. Luxembourg.
- OECD. (2014). *PISA 2012 results in focus: What 15-year-olds know and what they can do with what they know*. Paris
- OECD. (2016). *PISA 2015 results (Volume I). Excellence and equity in education*. Paris.
- Osborne, J. (2010). Arguing to learn in science: The role of collaborative, critical discourse. *Science*, 328(5977), 463-466.


- Osborne, J. (2013). The 21st century challenge for science education: Assessing scientific reasoning. *Thinking Skills and Creativity*, 10, 265-279.
- Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological Review*, 97, 185-200.
- Padilla, M. (1990). *The science process skills. Paper 9004 in the series, Research matters to the science teacher*. National Association for Research in Science Teaching (NARST). Retrieved May 6, 2016, from <https://www.narst.org/publications/research/skill.cfm>
- Pajares, F., & Schunk, D. H. (2001). Self-beliefs and school success: Self-efficacy, self-concept, and school achievement. In R. Riding & S. Rayner (Eds.), *BT self-perception* (pp. 239-66). London: Ablex Publishing.
- Partnership for 21st Century Skills. (2006). *A state leader's action guide to 21st century skills: A new vision for education*. Tucson, AZ: Partnership for 21st Century Skills.
- Pásztor, A. (2016). Az induktív gondolkodás technológia alapú mérése és fejlesztése. [Technology-based assessment and development of inductive reasoning]. Unpublished Doctoral dissertation, University of Szeged, Hungary). Retrieved from <http://doktori.bibl.u-szeged.hu/3191/>
- Pásztor, A., & Molnár, G. (2013). Validity and effectiveness of computer based versus face to face testing in first grade. *15th European Conference for the Research on Learning and Instruction* (pp. 27-31). Munich: EARLI Press.
- Pásztor, A., Molnár, G., & Csapó, B. (2015). Technology-based assessment of creativity in educational context: The case of divergent thinking and its relation to mathematical achievement, *Thinking skills and Creativity*, 18, 32-42.
- Pásztor, A., Molnár, G., Korom, E. B., Németh, M., & Csapó, B. (2017). Online assessment of inductive reasoning and its predictive power on inquiry skills in science. *17th biennial Conference of the European Association for Research on Learning and Instruction (EARLI)*. (p. 509), Tampere, Finland, August 29- September 2, 2017.
- Piaget, J. (1970). Piaget's theory. In P. H. Mussen (Ed.), *Carmichael's manual of child psychology* (Vol. 1). New York: Wiley.
- Piaget, J., & Inhelder, B. (1969). *The psychology of the child*. New York: Basic Books.
- Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology*, 95, 667-686.
- Piraksaa, C., Srisawasdi, N., & Koul, R. (2014). Effect of gender on students' scientific reasoning ability: A case study in Thailand, *Procedia - Social and Behavioral Sciences* 116, 486-491.
- Rambuda, A. M., & Fraser, W. F. (2004). Perceptions of teachers of the application of science process skills in the teaching of Geography in secondary schools in the Free State province. *South African Journal of Education*, 24, 10-17.
- Reddy, V., Zuze, T. L., Visser, M., Winnaar, L., Juan, A., Prinsloo, C., & Rogers, S. (2015). *Beyond bench marks: What twenty years of TIMSS data tell us about South-African education*. Cape Town: HSRC.

- Resing, W. C., Bakker, M., Pronk, C. M., & Elliott, J. G. (2016). Dynamic testing and transfer: An examination of children's problem-solving strategies. *Learning and Individual Differences*, 49, 110-119.
- Resing, W. C., Bakker, M., Pronk, C. M., & Elliott, J. G. (2017). Progression paths in children's problem solving: the influence of dynamic testing, initial variability, and working memory. *Journal of Experimental Child Psychology*, 153, 83-109.
- Roadranga, V., Yeany, R., & Padilla, M. J. (1982). *GALT, Group of Assessment of Logical Thinking* (pp 127-152). Athens, GA: University Of Georgia.
- Roberts, D. A. (2007). Scientific literacy/science literacy. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research in science education* (pp. 729 – 779). Mahwah, NJ: Erlbaum.
- Ropo, E. (1987). Skills for Learning: A Review of Studies on Inductive Reasoning. *Scandinavian Journal of Educational Research* 31, 31-39.
- Sandberg, E. H., & McCullough, M. B. (2010). The development of reasoning skills. In E. H. Sandberg & B. L. Spritz (Eds.), *A clinician's guide to normal cognitive development in childhood* (pp. 179-189). New York: Routledge
- Sandene, B., Horkay, N., Bennett, R., Allen, N., Braswell, J., Kaplan, B., & Oranje, A. (2005). *Online assessment in mathematics and writing: Reports from the NAEP Technology- Based Assessment Project, Research and Development Series* (NCES 2005-457). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office. Retrieved March 12, 2010, from http://nces.ed.gov/nationsreportcard/pdf/studies/2005457_1.pdf
- Schaeffer, G. A., Reese, C. M., Steffen, M., McKinley, R. L., & Mills, C. N. (1993). Field test of a computer-based GRE general test. *ETS Research Report #93-07*. Available from ERIC: ED 385 588.
- Schunk, D. H., Pintrich, P. R., & Meece, J. L. (2008). *Motivation in education* (3rd ed). Upper Saddle River, NJ: Pearson.
- Schwartz, R. S., Lederman, N., & Lederman, N. (2008). An instrument to assess views of scientific inquiry: The VOSI questionnaire. *International conference of the National Association for Research in Science Teaching*, March, 2008, Baltimore, MD.
- Senler, B. (2015). Middle school students' views of scientific inquiry: An international comparative study. *Science Education International*, 26(2), 166-179.
- Shayer, M., & Adey, P. (Eds.). (2002). *Learning intelligence: Cognitive acceleration across the curriculum from 5 to 15 years*. Milton Keynes: Open University Press.
- Shye, S. (1988). Inductive and deductive reasoning: A structural reanalysis of ability tests. *Journal of Applied Psychology*, 73, 308-311.
- Simasiku, L., Kasanda, C. & Smit, T. (2015). Teaching subjects matter through English as the Medium of instruction in the Namibian English second language classrooms. *European Scientific Journal Special edition*, 1, 18 - 28.
- Sloman, S. A. (1993). Feature-based induction. *Cognitive Psychology*, 25, 231-280.

- Snow, R. E., Kyllonen, P. C., & Marshalek, B. (1984). The topography of ability and learning correlations. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 47-103). Hillsdale, NJ: Erlbaum
- Spearman, C. (1923). *The nature of intelligence and the principles of cognition*. London: Macmillan.
- Sternberg, R. J. (1998). When will the milk spoil? Everyday induction in human intelligence. *Intelligence*, 25(3), 185-203.
- Sternberg, R. J., & Gardner, M. (1983). Unities in inductive reasoning. *Journal of Experimental Psychology: General*, 112, 80-116.
- Sweller, J. (2005). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 147-158). New York: Cambridge University Press.
- Taylor, C. S., & Nolen, S. B. (2008). *Classroom assessment: Supporting teaching and learning in real classrooms*. New Jersey: Pearson Merrill Prentice Hall.
- Tisher, R.P., & Dale, L.G. (1975). *Understanding in science test*. Victoria: Australian Council for Educational Research.
- Tobin, K., & Capie, W. (1981). The development and validation of a group test of logical thinking. *Educational and Psychological Measurement*, 41(2), 423-459.
- Treagust, D. F. (1995). Diagnostic assessment of students' science knowledge. *Learning science in the schools: research reforming practice*. S. M. Glynn & R. Duit (Eds.). (pp. 327-346). Mahwah, NJ: Lawrence Erlbaum Associates.
- Turiman, P., Omar, J., Daud, A., & Osman, K. (2012). Fostering the 21st century skills through scientific literacy and science process skills. *Procedia-Social and Behavioral Sciences*, 59, 110-116.
- Tytler, R. (2007). *Re-imagining science education: Engaging students in science for Australia's future*. Camberwell: Vic: ACER Press.
- U.S. Department of Education. (2010). *National educational technology plan 2010 draft*. Office of Educational Technology. Retrieved from <http://www.ed.gov/technology/netp-2010>
- Upmeyer zu Belzen, A., & Kruger, D. (2010). "Modellkompetenz im Biologieunterricht" [Model Competence in Biology Lessons]." *Zeitschrift für Didaktik der Naturwissenschaften* 16: 41-57.
- Vainikainen, M. P. (2014). *Finnish primary school pupils' performance in learning to learn assessments: A longitudinal perspective on educational equity*. Unpublished doctoral thesis, University of Helsinki, Helsinki, Finland.
- Vainikainen, M. P., Hautamäki, J., Hotulainen, R., & Kupiainen, S. (2015). General and specific thinking skills and schooling: Preparing the mind to new learning. *Thinking Skills and Creativity*, 18, 53-64.
- Valamides, N. C. (1996). Formal reasoning and science teaching. *School Science and Mathematics*, 96(2), 99-107.
- Van de Vijver, F. J. R. (1991). *Inductive thinking across cultures: An empirical investigation*. Tilburg: Catholic Univ. of Brabant.

- Vygotsky, L. S. (1978). *Mind and Society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wainer, H., & Mislevy, R. J. (2000). Item response theory, calibration, and estimation. In H. Wainer (Ed.), *Computerized adaptive testing: A primer*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Weiss, D. J., & Kingsbury, G. G. (1984). Application of computerized adaptive testing to educational problems. *Journal of Educational Measurement*, 21, 361-375.
- Wenning, C. J. (2005a). Levels of inquiry: Hierarchies of pedagogical practices and inquiry processes. *Journal of Physics Teacher Education Online*, 2(3), 3-11.
- Wenning, C. J. (2006). Assessing nature of science literacy as one component of scientific literacy. *Journal of Physics Teacher Education Online*, 3-14.
- Wenning, C. J. (2007). Assessing inquiry skills as a component of scientific literacy. *Journal of Physics Teacher Education Online*, 4(2), 21-24.
- Wong, S. L., & Hodson, D. (2008). From the horse's mouth: What scientists say about scientific investigation and scientific knowledge. *Science Education*, 93, 109-130.
- Yeh, Y., Jen, T., & Hsu, Y. (2012). Major strands in scientific inquiry through cluster analysis of research abstracts. *International Journal of Science Education*, 34, 2811-2842.
- Zeineddin, A., & Abd-El-Khalick, F. (2010). Scientific Reasoning and Epistemological Commitments: Coordination of Theory and Evidence Among College Science Students. *Journal of research in science teaching*, 47(9), 1064-1093. doi:10.1002/tea.20368
- Zhou, S., Han, J., Koenig, K., Raplinger, A., Pi, Y., Li, D., & Bao, L. (2016). Assessment of scientific reasoning: Effect of task context, data and design on student reasoning in control of variable. *Thinking Skills and Creativity*, 19, 175-187.
- Zimmerman, C. (2005). *The development of scientific reasoning skills: What psychologists contribute to an understanding of elementary science learning?* National Research Council.
- Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review*, 27, 172-223.

Appendix A: Scientific reasoning test (5th and 7th grades)







Scientific Reasoning Assessment
MTA-SZTE Research Group on The Development Competences
H-6722 Szeged, Petőfi Sándor egt. 30-34. Hungary





Scientific reasoning test

Please enter your code and then click on the [Next](#) button.

Your code:


[Enter](#)





Scientific Reasoning Assessment
MTA-SZTE Research Group on The Development Competences
H-6722 Szeged, Petőfi Sándor egt. 30-34. Hungary

Scientific reasoning test



Dear student!

Welcome to our test! We will examine your way of thinking. You can expect interesting tasks. Read the instructions carefully before you give your answer.


If you find a task too difficult you can click on the [Next](#) button and go on. Later you can return to the task by clicking on the [Previous](#) button.

The [orange line](#) at the top of the screen indicates your current position in the test. After the last task you will see your result.

You have [40 minutes](#) to complete the test.

Click on the [Next](#) button.


[Next](#)



Are you a girl or a boy? Click on the answer.


☐ Boy

☐ Girl



Next

We pour the milk from the glass into the bowl. Which statement is true?
Click on the answer.



☐ Both the volume and the shape of the milk change.

☐ Only the volume changes, but not the shape of the milk.

☐ Only the shape of the milk changes, but not the volume.

☐ Neither the shape nor the volume changes.

Previous

Next

We move the marble from a smaller glass into the bigger one. Which statement is true?
Click on the answer.



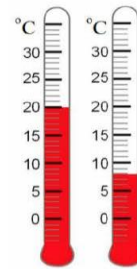
- ☐ It changes both the volume and the shape of the marble.
- ☐ It changes only the volume but not the shape of the marble.
- ☐ It changes only the shape but not the volume of the marble.
- ☐ It does not change either the shape or the volume of the marble.

[Previous](#)

[Next](#)

Maria wanted to know what the temperature outside was in the morning. She took the thermometer from the room to outside. The picture shows the change that occurred after a few minutes. Which properties of the thermometer liquid changed? Click on the answer.

- | | | |
|---------|--------------------------------|------------------------------------|
| Mass | <input type="radio"/> Changed. | <input type="radio"/> Not changed. |
| Volume | <input type="radio"/> Changed. | <input type="radio"/> Not changed. |
| Density | <input type="radio"/> Changed. | <input type="radio"/> Not changed. |
| Shape | <input type="radio"/> Changed. | <input type="radio"/> Not changed. |



[Previous](#)

[Next](#)

Which of the following changes is reversible? Click on the answer.

We burn the firewood.

☐ Yes

☐ No

We dissolve sugar in water.

☐ Yes

☐ No

We slice up the bread.

☐ Yes

☐ No

We warm up the water.

☐ Yes

☐ No

[Previous](#)

[Next](#)

Does iron element under go substance of the iron change in the following situations?
Click on the answer.

it is rusting

☐ Yes

☐ No

it is bending

☐ Yes

☐ No

it is placed near to a magnet

☐ Yes

☐ No

it is falling off

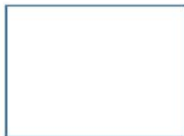
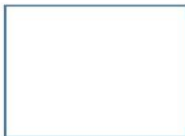
☐ Yes

☐ No

[Previous](#)

[Next](#)

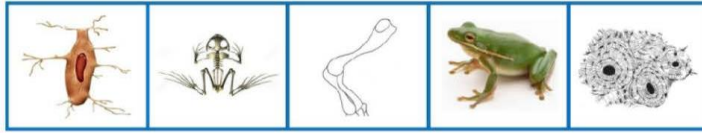
Put the pictures in the right order. Pull them to the boxes.



[Previous](#)

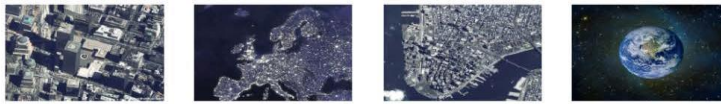
[Next](#)

Put the pictures in the right order. Pull them to the boxes.



--	--	--	--	--

Put the pictures in the right order. Pull them to the boxes.



--	--	--	--

Previous

Next

You have to change two pictures to get the right order. Which are those two pictures?
Click on them.



Previous

Next

Which is the odd one out? Click on the answer.

- ☐ fruiting body
- ☐ ring
- ☐ spore
- ☐ root
- ☐ stalk

[Previous](#)

[Next](#)






Which one of the four statements **does not** belong in this group.
Click on the answer.

- ☐ The puddle dries.
- ☐ The branch gets frosty.
- ☐ The river floods.
- ☐ The dew falls on the handrail.

[Previous](#)

[Next](#)

Which picture **does not** belong in this group? Click on it.



[Previous](#) [Next](#)

Look at the connection between the two words in the blue boxes.
Which word should take the place of the question mark? Click on it.

metal	plastic
-------	---------

solid	?
-------	---

[iron](#) [liquid](#) [state of matter](#) [wood](#)

[Previous](#) [Next](#)

Look at the connection between the two words in the blue boxes.
Which word should take the place of the question mark? Click on it.

Sun	Earth
-----	-------

magnet	?
--------	---

[wooden spoon](#) [needle](#) [ceramic knife](#) [calculator](#)

[Previous](#) [Next](#)

Observe the common characteristics of the pictures that are placed side by side and under each other. Which picture should be replaced with the separate (picture on the right) one? Click on it.



[Previous](#)

[Next](#)

Observe the common characteristics of the pictures that are placed side by side and under each other. Which picture should be replaced with the separate one? Click on it.



[Previous](#)

[Next](#)

Observe the common characteristics of the substances that are placed side by side and under each other. Which word should be replaced with the separate one? Click on it.

carbon-dioxide	cooking oil	sugar
air	vinegar	gravel

oxygen

[Previous](#)

[Next](#)

Mathew, Rosaria and Ndahafa, marked cities as travel destinations at a distance of 10 cm from the capital city on maps using different scales. Which city should be best approached by bicycle, car or plane? Choose the appropriate means of transport for each child. Click on the answers.

Students	Scale on the map	Most practical means of transport		
Matthew	1 : 1 500 000	<input type="radio"/> bicycle	<input type="radio"/> car	<input type="radio"/> plane
Rosaria	1 : 40 000	<input type="radio"/> bicycle	<input type="radio"/> car	<input type="radio"/> plane
Ndahafa	1 : 11 600 000	<input type="radio"/> bicycle	<input type="radio"/> car	<input type="radio"/> plane

[Previous](#)

[Next](#)

A kangaroo can move two times faster than an elephant. Which statement is true? Click on the answer.



A kangaroo can do twice more distance than an elephant in the same amount of time.

☐ True

☐ False



A kangaroo can do the same amount of distance in twice more time than an elephant.

☐ True

☐ False

[Previous](#)

[Next](#)



The rotation period of Earth is one day, namely 24 hours. During this time it rotates 360° around its own axis. Choose the belonging day/hour/angle amounts. Click on the answers.

The time of rotation (day)	The time of rotation (hour)	The angle of rotation
1,5	<input type="text" value="Choose!"/>	<input type="text" value="Choose!"/>
<input type="text" value="Choose!"/>	8	<input type="text" value="Choose!"/>

[Previous](#)

[Next](#)

Learners boiled 2 litres of water to 80°C in a pot. The temperature of water was 20°C at the beginning and it took 5 minutes to boil the water. How much time will they need to boil 2 litres of water to 80°C , if the temperature of water is 50°C at the beginning? Click on the answer.



• 2.5 minutes

• 5 minutes

• 7 minutes

• 10 minutes

◀ Previous

Next ▶

How many times can you visit your grandmother, who lives 10 km away, if you have 3 litres of petrol and your motorcycle consume 5 litres of petrol per 100 km? Click on the answer.



• twice

• three times

• four times

• six times

◀ Previous

Next ▶

The student had a medical examination in school. It was found that some of the children in the classes were overweight. The following table shows the data for the three classes. Does being overweight depend on whether the child is a boy or a girl? Click on the answer.

Gender	Number of students	
	Owerweight	Normal weight
Boys	8	38
Girls	11	43

• Yes

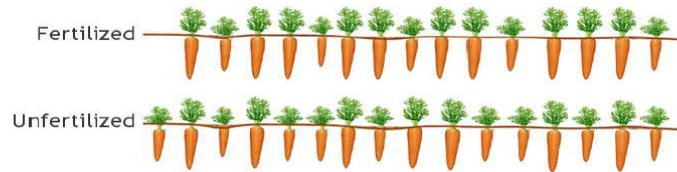
• No

• We can't decide.

◀ Previous

Next ▶

Did fertilization influence the size of the carrots? Click on the answer.



- ☐ Yes ☐ No ☐ We can't decide.

[Previous](#)

[Next](#)

There are events which will definitely take place, and there are events that may not take place. Decide which group the following events belong to. Click on the answer.

- The house collapse during an earthquake. ☐ Certain ☐ Uncertain
- If one is born, one will die someday. ☐ Certain ☐ Uncertain
- It is snowing at Christmas. ☐ Certain ☐ Uncertain
- A flower turns to fruit. ☐ Certain ☐ Uncertain

[Previous](#)

[Next](#)

Look at the pictures below.
Give the four pictures a common title that shows their similarity. Click on the answer.

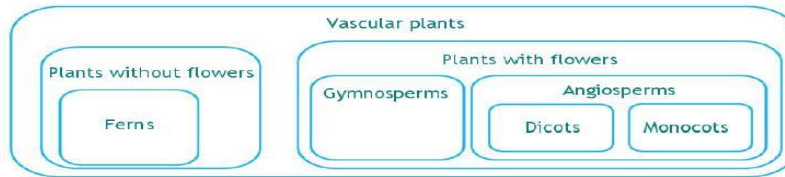


- ☐ Winter ☐ Water ☐ Weather ☐ Ice

[Previous](#)

[Next](#)

Observe the figure. Decide whether the following statements are true or false.
Click on the answer.



If a plant is a dicot, it will not have flowers.

☐ True ☐ False

Gymnosperms are vascular plants.

☐ True ☐ False

There are some flowering plants which are monocots.

☐ True ☐ False

Every ferns are flowerless.

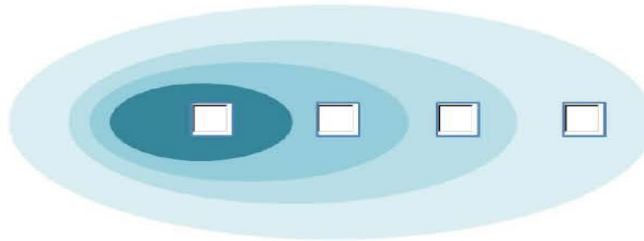
☐ True ☐ False

[Previous](#)

[Next](#)

Write the numbers of the expressions to the right place on the figure.

1 carpel 2 plant 3 ovule 4 flower

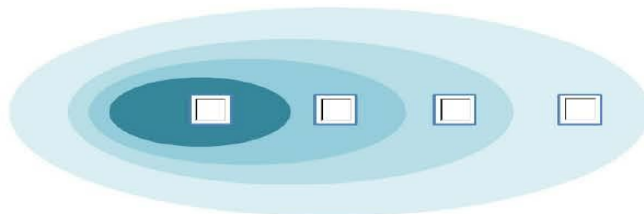


[Previous](#)

[Next](#)

Write the numbers of the expressions to the right place on the figure.

1 anthracite 2 substances 3 coals 4 fossil fuels



[Previous](#)

[Next](#)



Scientific reasoning test



You have reached the **END** of the test.

If you click on the **Next** button you will not be able to step back to the tasks.

Click on the **Next** button to see your result.

[Previous](#)

[Next](#)

Appendix B: Inductive reasoning test

Inductive reasoning test

Dear students, welcome to our Inductive reasoning skills test! The purpose of this test is to examine your way of **thinking and reasoning skills**. This is a test of your ability to apply aspects of inductive reasoning, analyze a situation and to make a prediction or to solve a problem.

INSTRUCTIONS

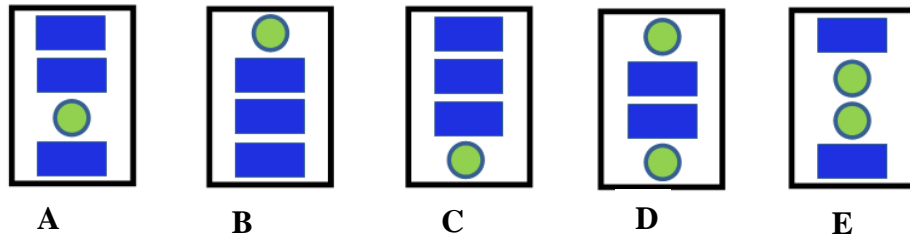
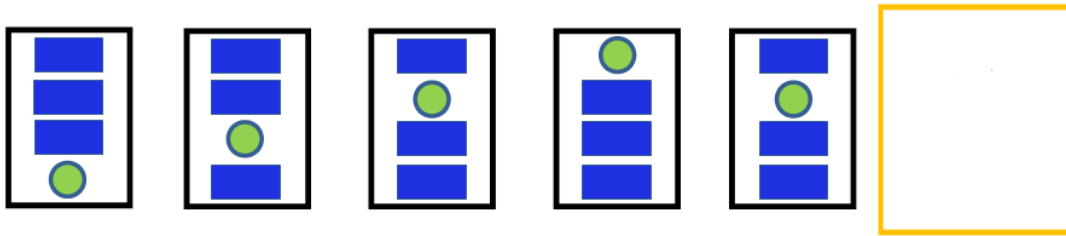
- Answer on the separate answer sheet provided
- Write your candidate number on top of each sheet you have used
- Write neatly and legible.
- Number your answers accordingly.
- For multiple choice questions, choose the letter that has the correct answer.
- **NB:** Read the instructions carefully before you answer each question.
- **NB: DO NO WRITE ANYTHING ON THE QUESTION PAPER BOOKLET!!!**
- You have **2H45 minutes** to complete the test.

Marks: 66

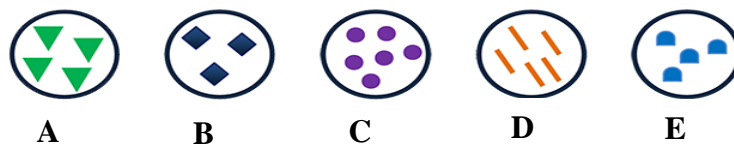
SECTION A:

Find the rules and continue the line!

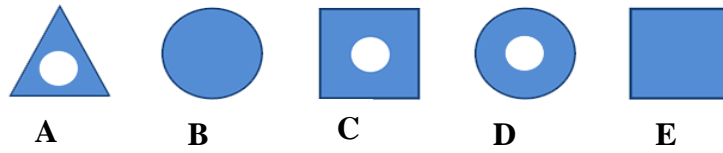
1. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)



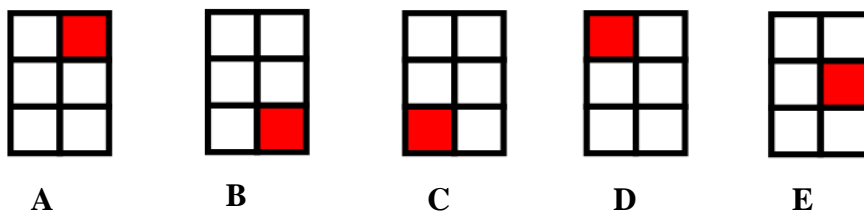
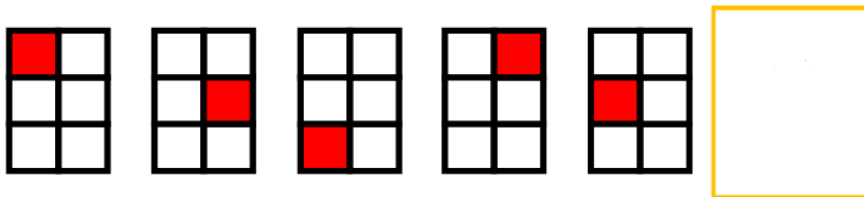
2. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)



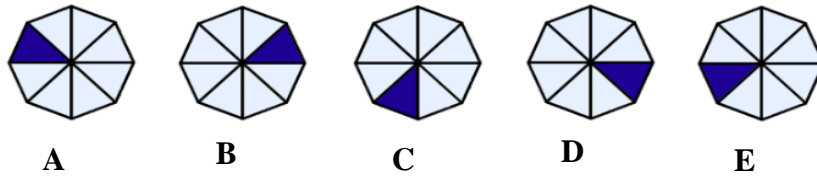
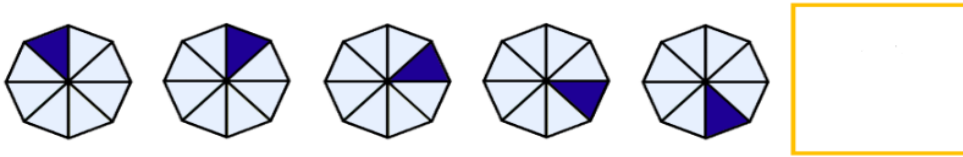
3. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)



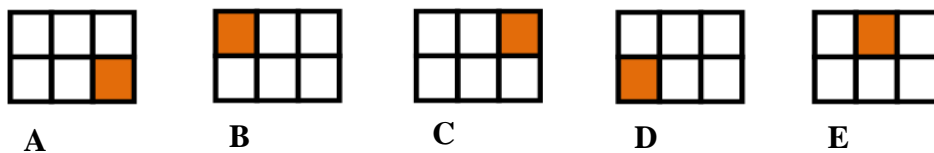
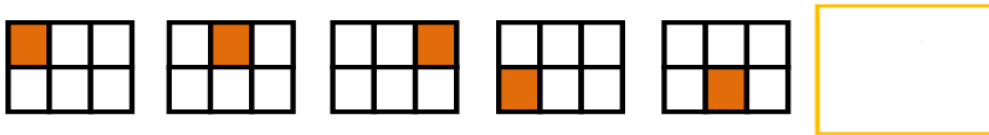
4. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)



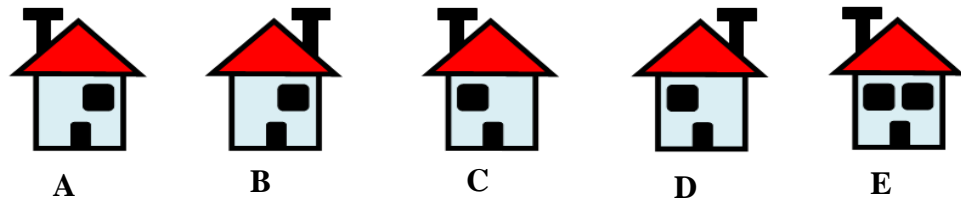
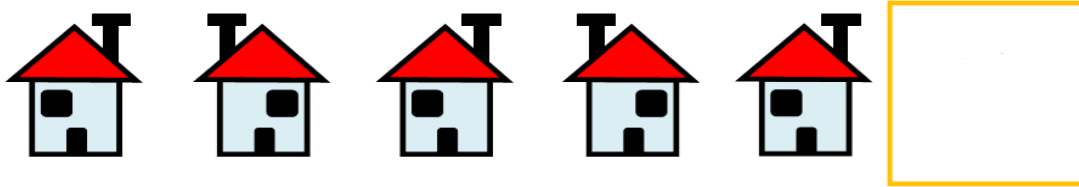
5. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)



6. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)



7. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)



8. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)



9. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)



A



B



C

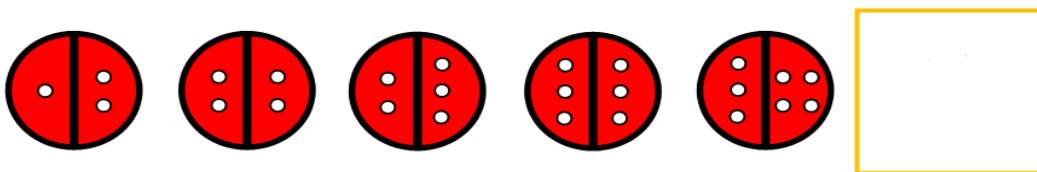


D



E

10. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)



A



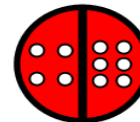
B



C

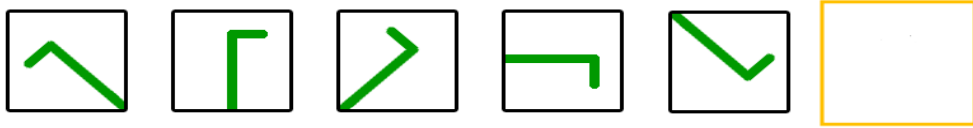


D



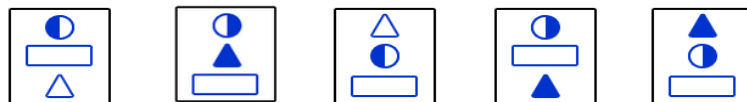
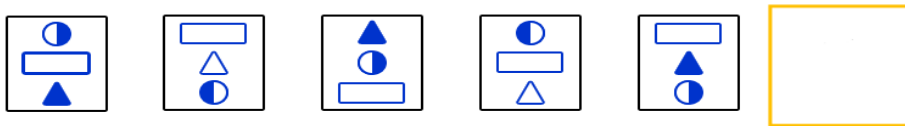
E

11. Which picture fits the most to the **yellow frame**? Write down the correct letter! (1)



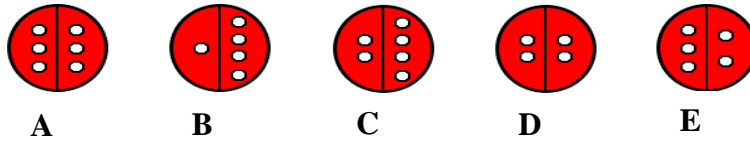
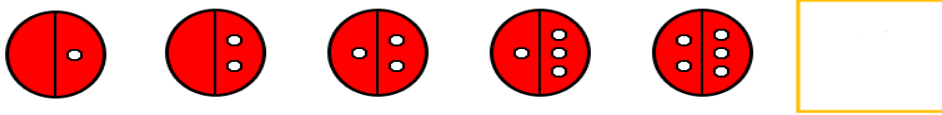
A B C D E

12. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)

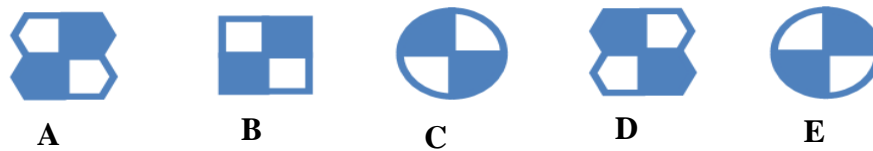


A B C D E

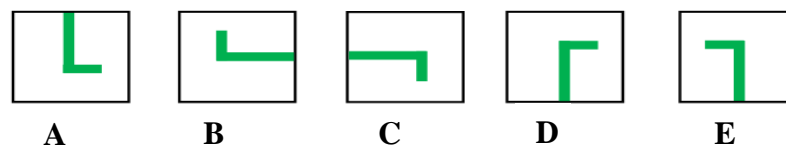
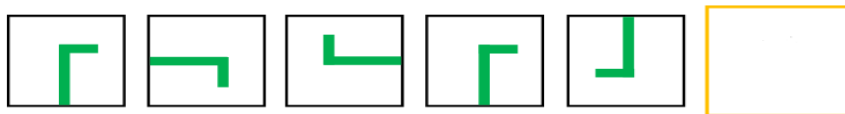
13. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)



14. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)



15. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)

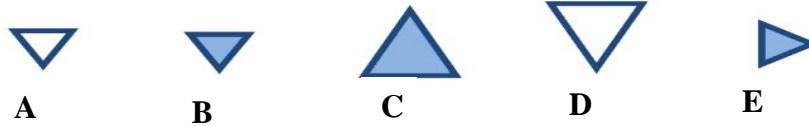
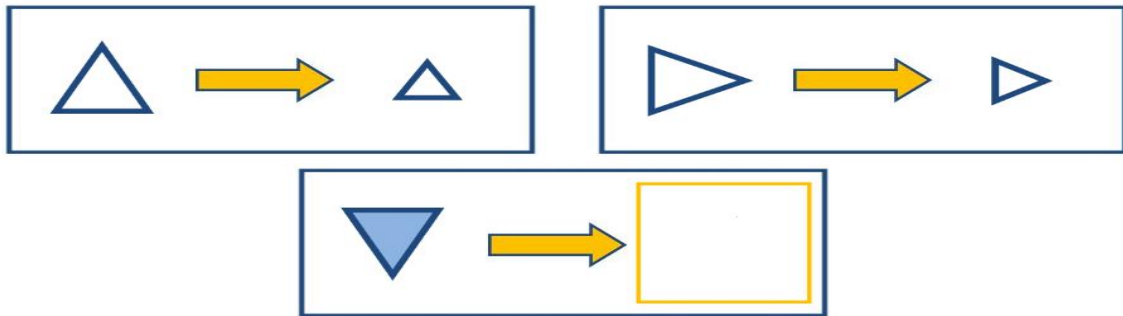


Total for section A: (15 Marks)

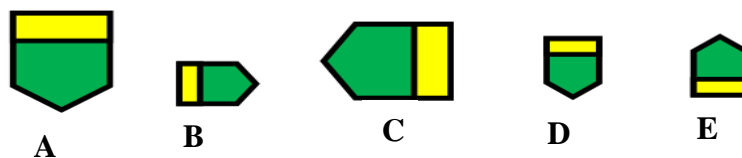
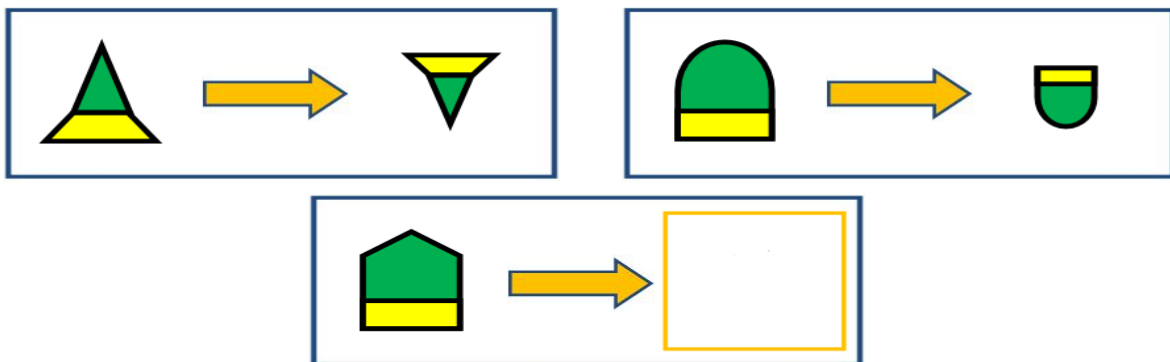
SECTION B:

In this section, examine how the pictures change in the blue frames! What could be the rule?

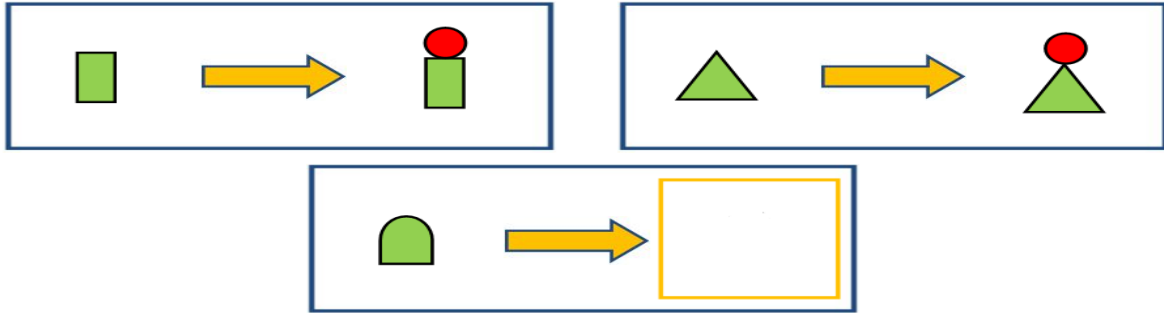
16. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)








17. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)

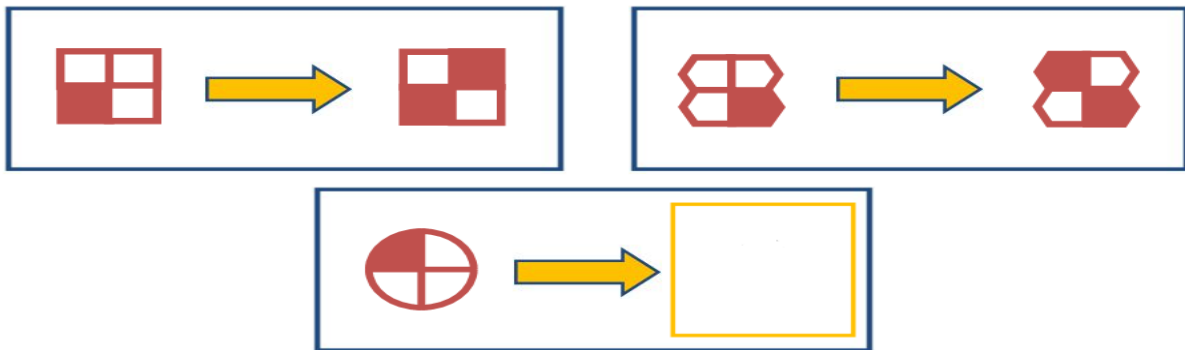







18. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)



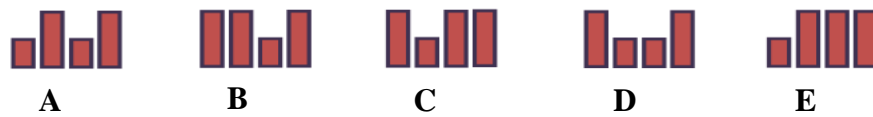
- 




- A B C D E

19. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)

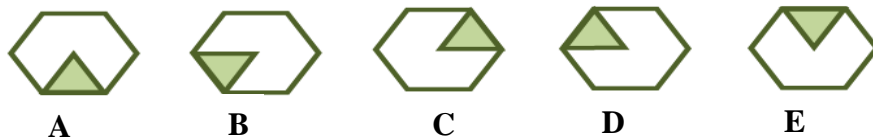
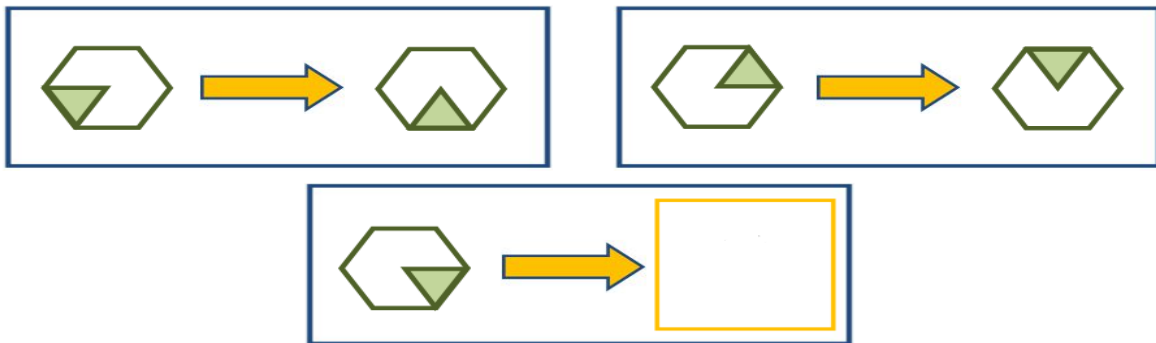


- 




- A B C D E

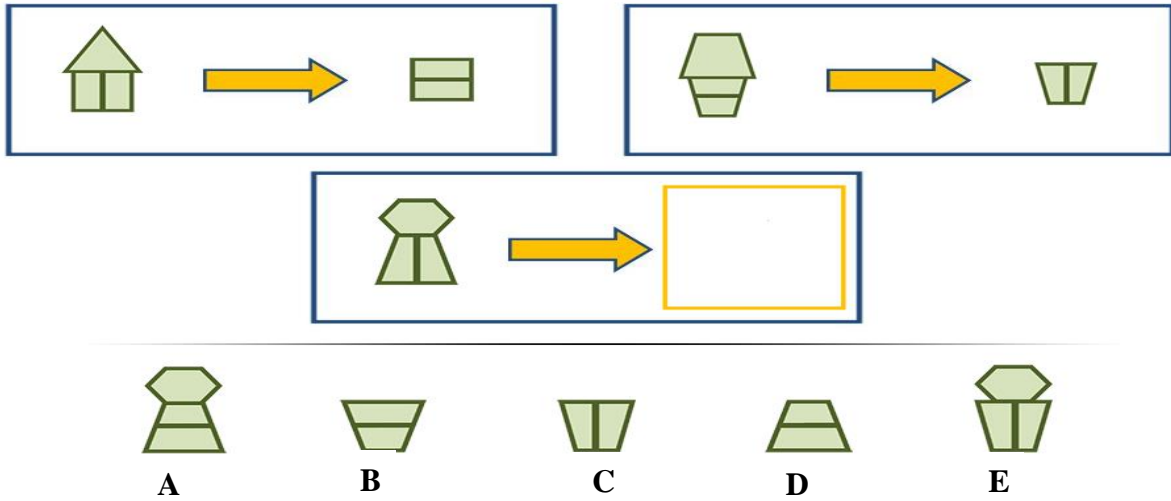
20. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)



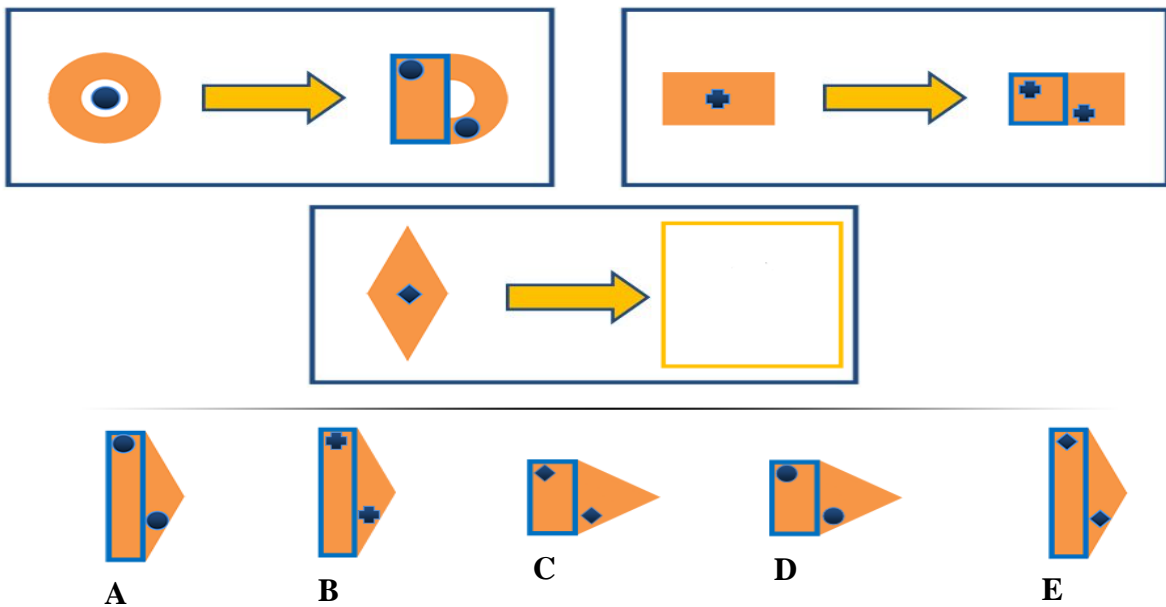
21. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)



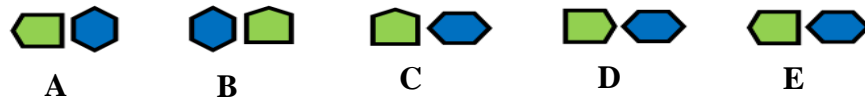
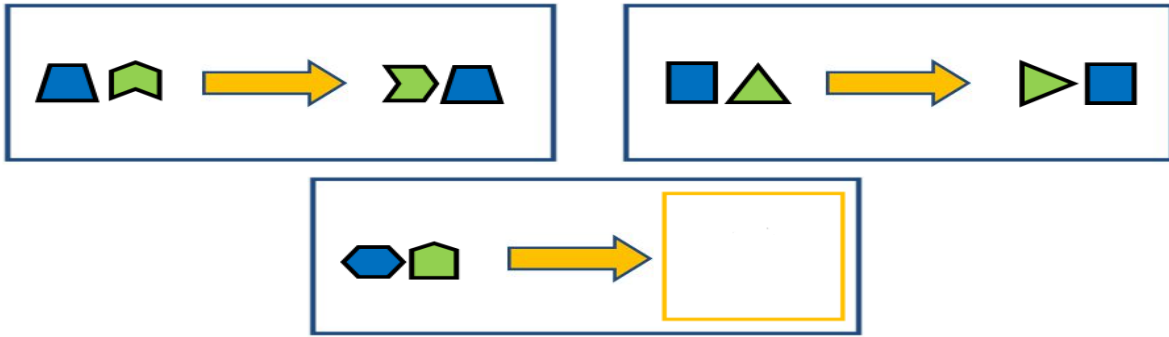
22. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)



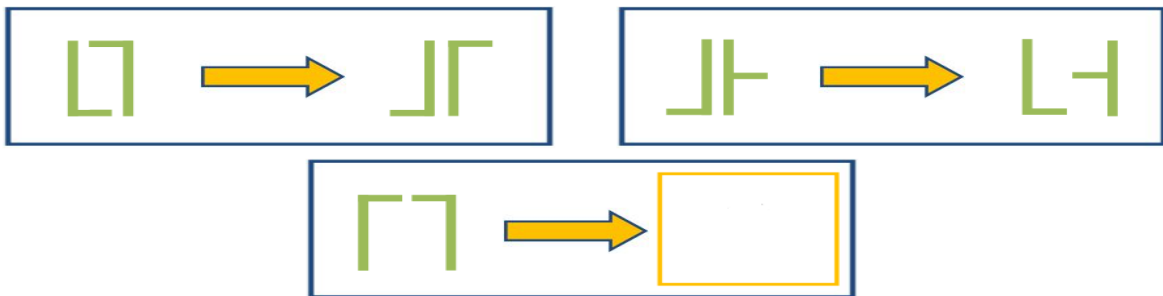
23. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)



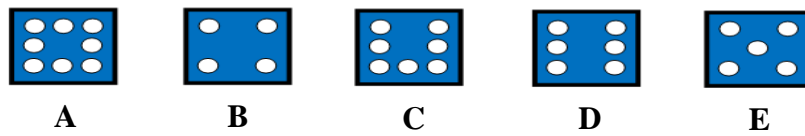
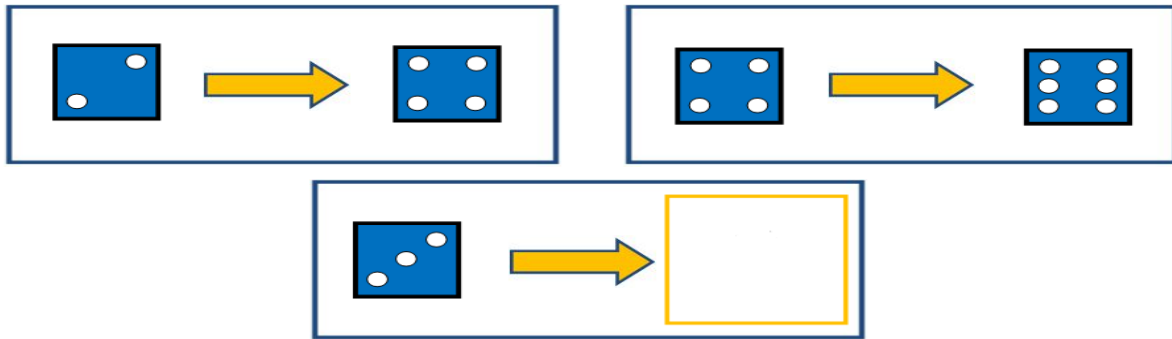
24. Which picture fits the most into the **yellow frame**? Write down the correct! (1)



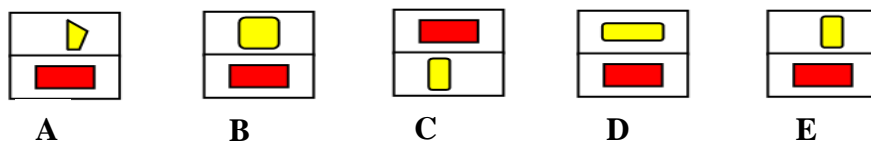
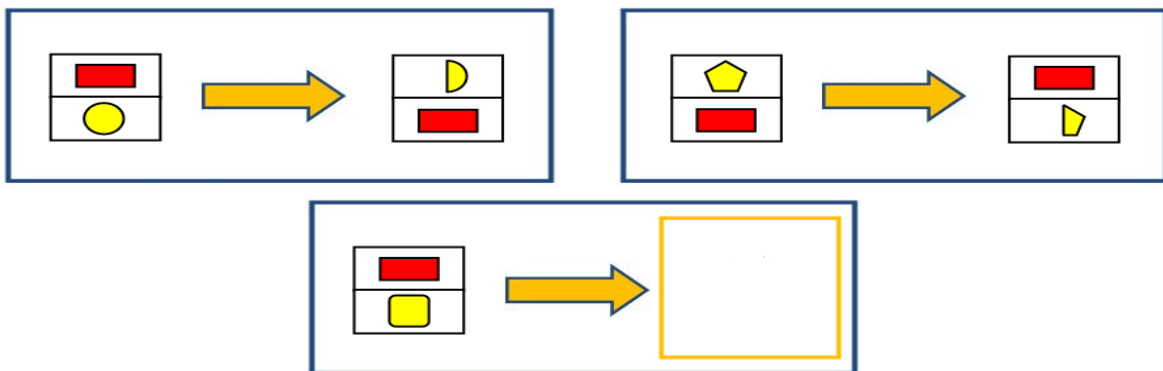
25. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)



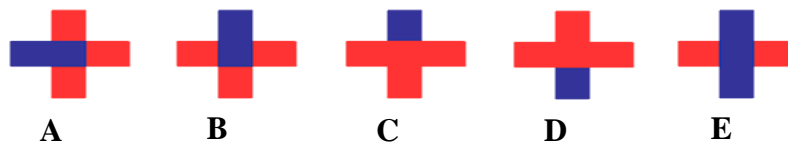
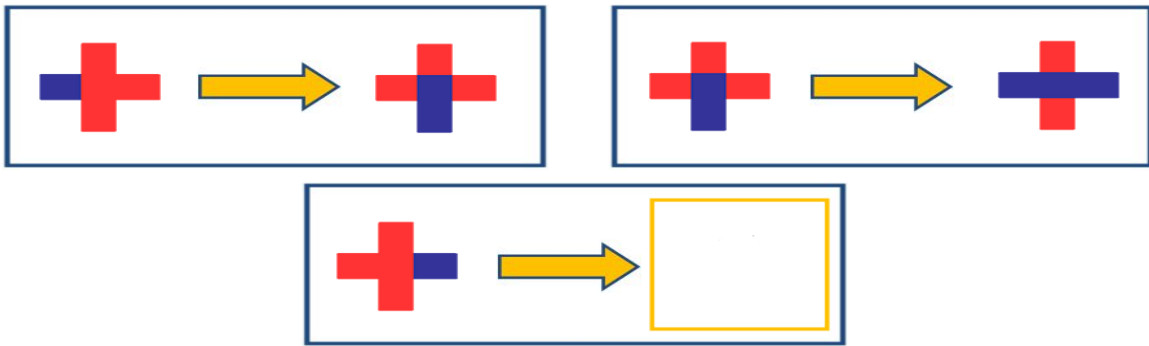
26. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)



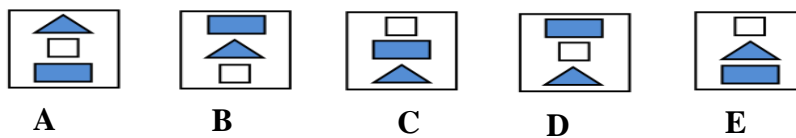
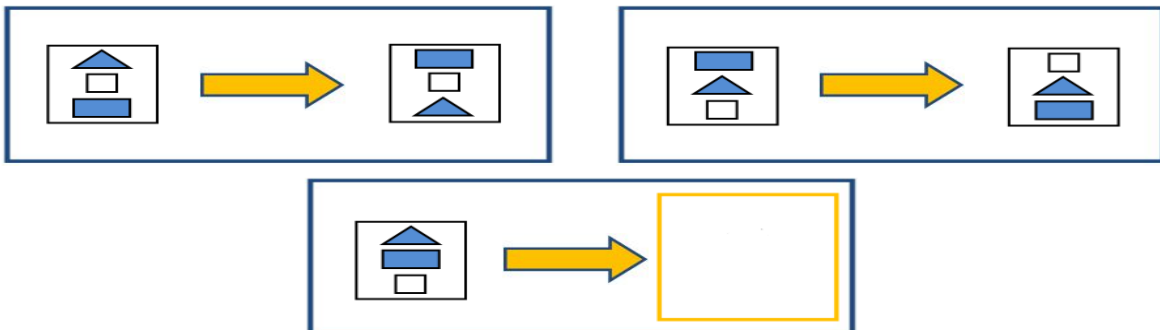
27. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)



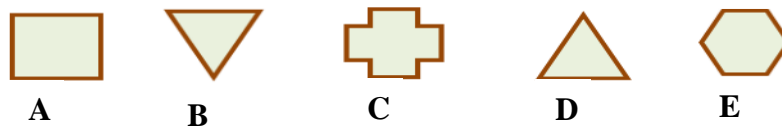
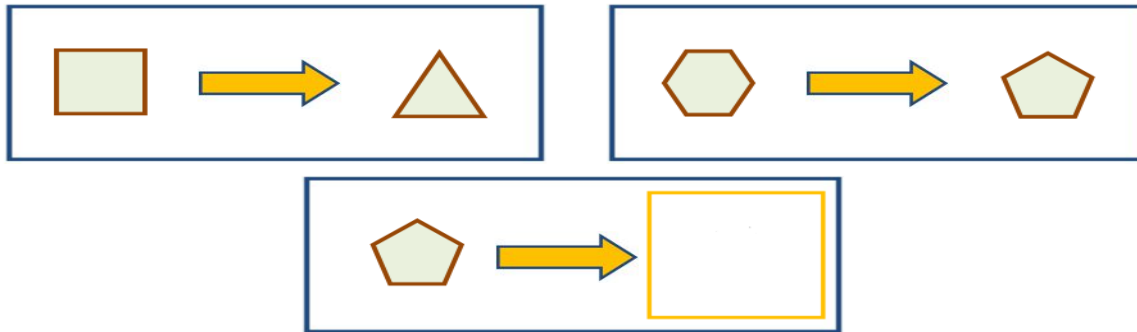
28. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)



29. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)



30. Which picture fits the most into the **yellow frame**? Write down the correct letter! (1)



Total for Section B: (15 Marks)

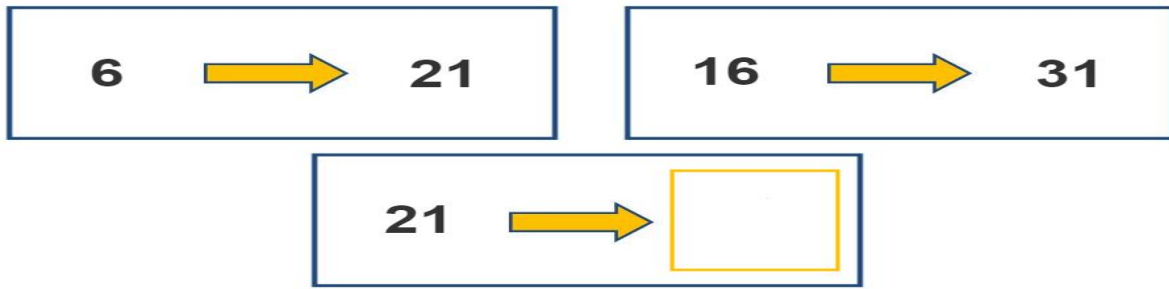
SECTION C:

In each case find the rule and choose the appropriate number that fits in the yellow box!

31. Which number fits the most into the **yellow frame**? Write down the correct letter! (1)



32. Which number fits the most into the **yellow frame**? Write down the correct letter! (1)



31

A

41

B

6

C

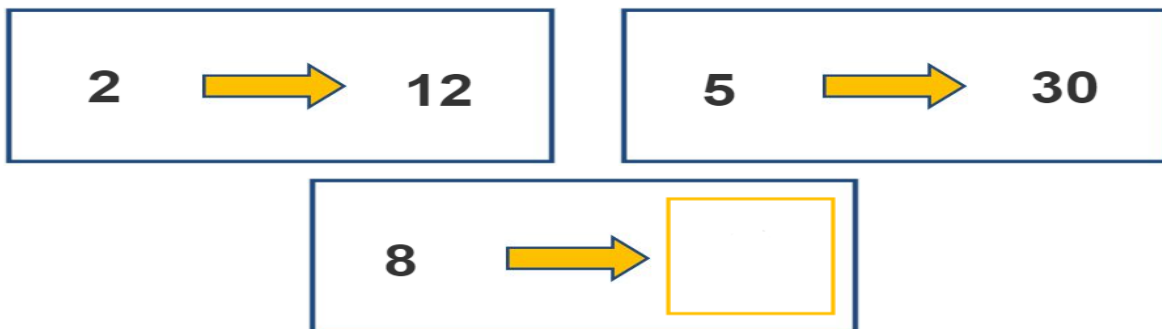
15

D

36

E

33. Which number fits the most into the **yellow frame**? Write down the correct letter! (1)



28

A

18

B

48

C

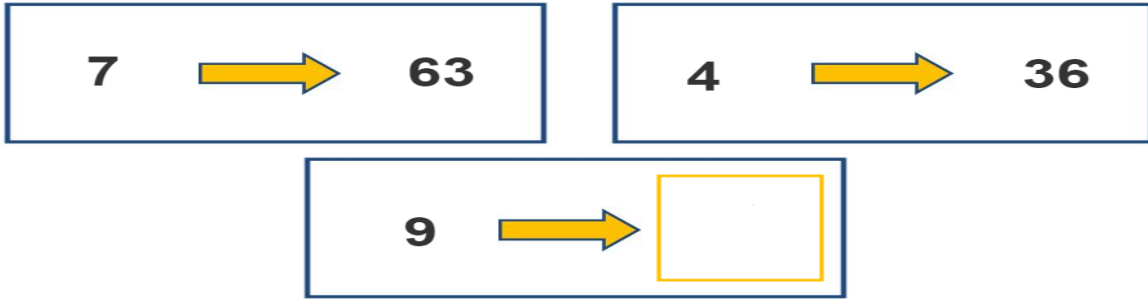
40

D

38

E

34. Which number fits the most into the **yellow frame**? Write down the correct letter! (1)



65
A

81
B

41
C

36
D

63
E

35. Which number fits the most into the **yellow frame**? Write down the correct letter! (1)



36
A

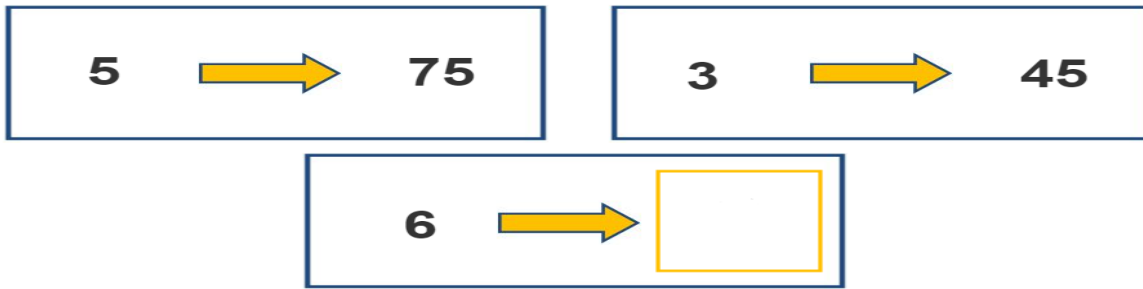
74
B

24
C

50
D

26
E

36. Which number fits the most into the **yellow frame**? Write down correct letter! (1)



90

A

76

B

80

C

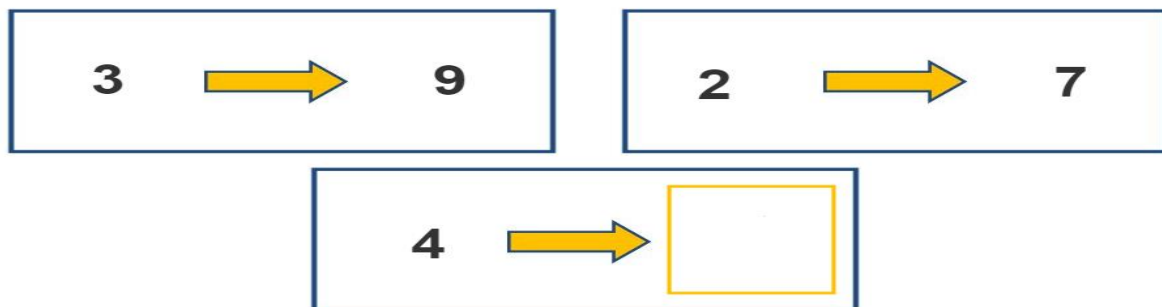
78

D

30

E

37. Which number fits the most into the **yellow frame**? Write down the correct letter! (1)



10

A

14

B

12

C

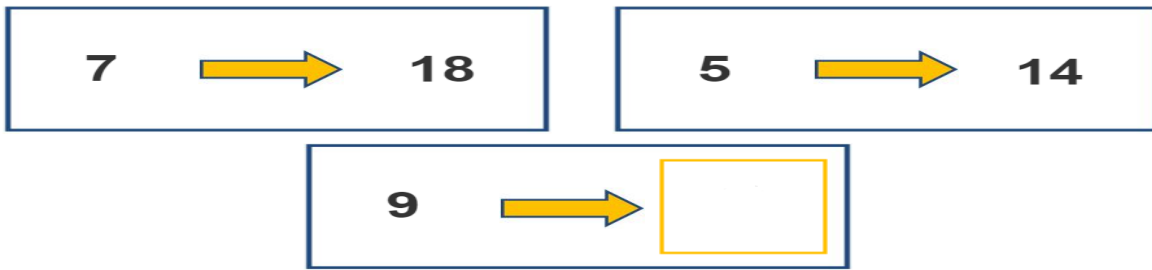
11

D

8

E

38. Which number fits the most into the **yellow frame**? Write down the correct letter! (1)



16

A

22

B

20

C

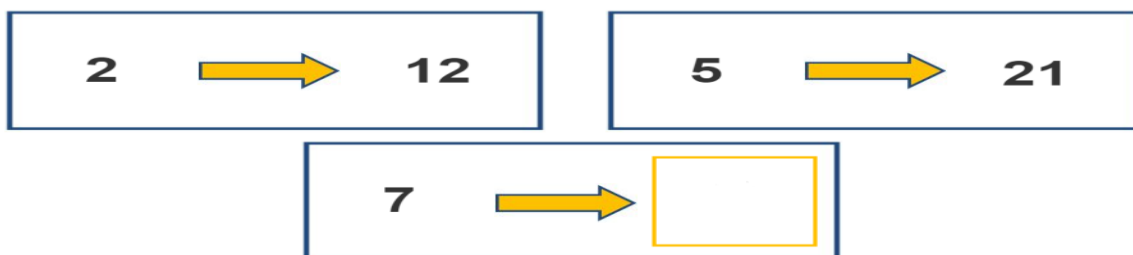
18

D

10

E

39. Which number fits the most into the **yellow frame**? Write down the correct letter! (1)



27

A

29

B

17

C

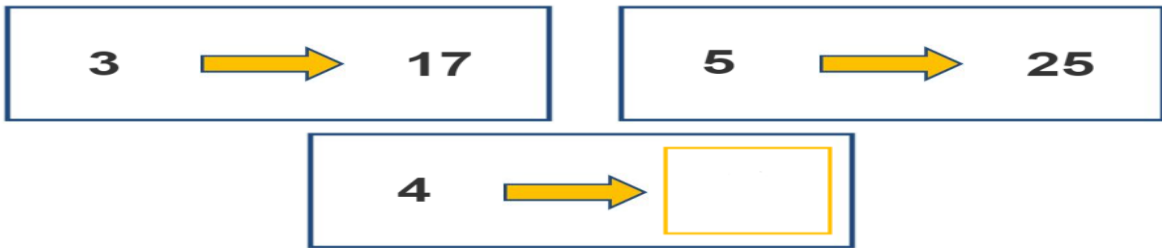
30

D

33

E

40. Which number fits the most into the **yellow frame**? Write down the correct letter! (1)



30

A

18

B

21

C

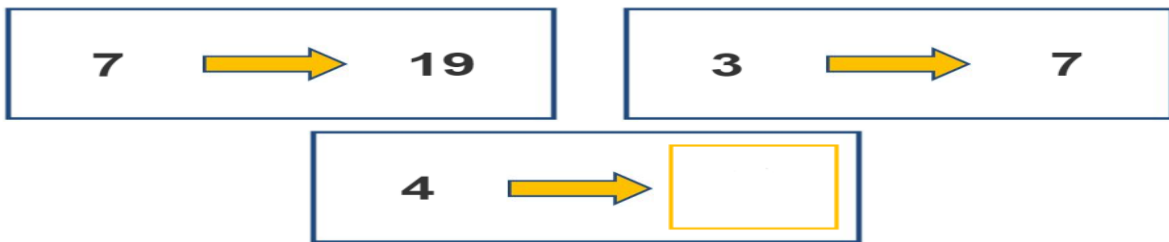
20

D

36

E

41. Which number fits the most into the **yellow frame**? Write down the correct letter! (1)



9

A

8

B

10

C

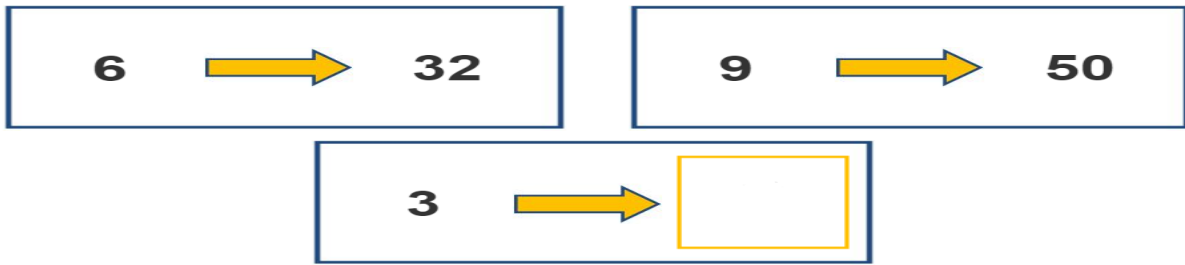
12

D

16

E

42. Which number fits the most into the **yellow frame**? Write down the correct letter! (1)



18
A

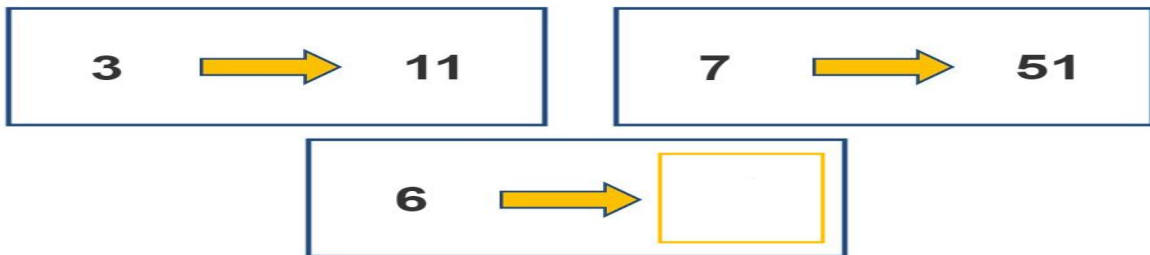
14
B

29
C

16
D

23
E

43. Which number fits the most into the **yellow frame**? Write down the correct letter! (1)



41
A

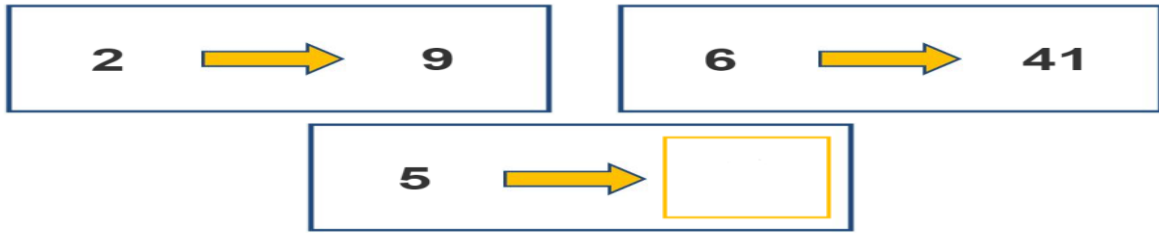
14
B

42
C

91
D

38
E

44. Which number fits the most into the **yellow frame**? Write down the correct letter! (1)



- 12** **33** **35** **30** **40**
A B C D E

Total for section C: (14 Marks)

SECTION D:

In this section, two numbers are needed to continue the number line!

45. Two numbers need to go into the **yellow frames** in order to **complete the number line best!**
Write down the correct letters. (2)



- 33** **48** **40** **64** **35** **128** **124**
A B C D E F G

46. Two numbers need to go into the **yellow frames** in order to **complete the number line best!**
Write down the correct letters! (2)



10 **13** **8** **11** **9** **12** **14**
A B C D E F F

47. Two numbers need to go into the **yellow frames** in order to **complete the number line best!** Write down the correct letters! (2)



26 **28** **30** **25** **32** **27** **29**
A B C D E F G

48. Two numbers need to go into the **yellow frames** in order to **complete the number line best!**
Write down the correct letters! (2)



26 **17** **20** **19** **22** **21** **29**
A B C D E F G

49. Two numbers need to go into the **yellow frames** in order to **complete the number line best.** Write down the correct letters! (2)



14 **21** **18** **34** **19** **23** **26**
A B C D E F G

50. Two numbers need to go into the **yellow frames** in order to **complete the number line best!**
Write down the correct letters! (2)



16
A

25
B

31
C

41
D

27
E

19
F

15
G

51. Two numbers need to go into the **yellow frames** in order to **complete the number line best!**
Write down the correct letters! (2)



99
A

61
B

109
C

257
D

129
E

57
F

217
G

52. Two numbers need to go into the **yellow frames** in order to **complete the number line the best!** Write down the correct letters! (2)



31
A

32
B

24
C

17
D

27
E

42
F

33
G

53. Two numbers need to go into the **yellow frames** in order to **complete the number line best!** Write down the correct letters! (2)



145
A

161
B

146
C

162
D

151
E

200
F

281
G

54. Two numbers need to go into the **yellow frames** in order to **complete the number line best!**
Write down the correct letters! (2)



109	84	128	94	83	90	190
A	B	C	D	E	F	G

55. Two numbers need to go into the **yellow frames** in order to **complete the number line best!** Write down the correct letters! (2)



29	42	64	37	48	27	45
A	B	C	D	E	F	G

Total for section D: (22 Marks)

You have reached **the end of the test**. Thank you so much for your participation!

Appendix C: Scientific inquiry skills test

(9th, 10th, 11th & 12th grades)

Dear students, welcome to our Scientific inquiry skills test! The purpose of this test is to examine your way of **thinking and reasoning skills**. This is a test of your ability to apply aspects of scientific reasoning skills, analyze a situation and to make a prediction or to solve a problem.

INSTRUCTIONS

- Answer on the separate answer sheet provided
- Write your candidate number on top of each sheet you have used
- Write neatly and legible.
- Number your answers accordingly.
- For multiple choice questions, choose the letter that has the correct answer.
- **NB:** Read the instructions carefully before you answer each question.
- **NB: DO NO WRITE ANYTHING ON THE QUESTION PAPER BOOKLET!!!**
- You have **2H45 minutes** to complete the test.

Background information

1. How old are you?

2. Are you a boy or a girl?

A) boy

B) girl

3. In which grade are you?

A) Grade 10

B) Grade 12

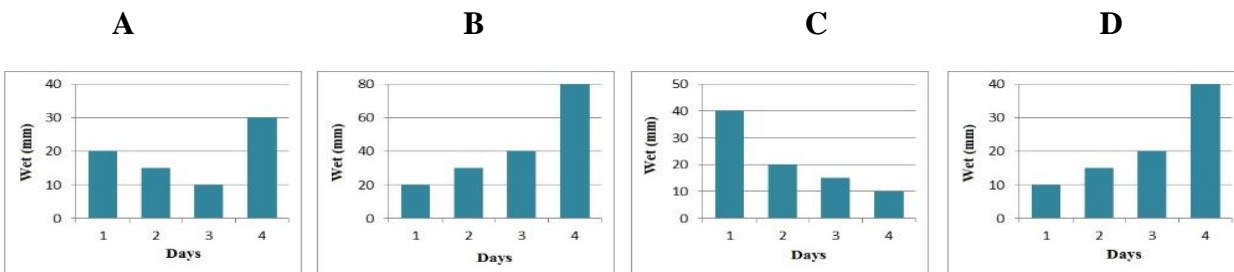
4. Where do you live? Choose one.

A) Town

B) Village

Section A

1. It was raining a lot in the previous days. On the first day 10 mm of rain fell, on the second day 15 mm, on the third day 20 mm and on the fourth day 40 mm of rain fell. Which diagram shows correctly represent rainfall in the past days? (1)



2. Ben observed the effect of exercise on the body. After running 500 meters, he measured his pulses every two minutes. Immediately after running, his pulse was 150, after two minutes it was 120, after 4 minutes it was 100, after 6 minutes it was 94 and after 8 minutes, it was 80. He recorded his measurements in a table.

Which table shows correctly his measurements?

(1)

A	
Time (minute)	Pulse/minute
0	80
2	94
4	100
6	120
8	150

B	
Time (minute)	Pulse/minute
2	80
4	94
6	100
8	120
10	150

C	
Time (minute)	Pulse/minute
0	150
2	120
4	100
6	94
8	80

D	
Time (minute)	Pulse/minute
8	150
6	120
4	100
2	94
0	80

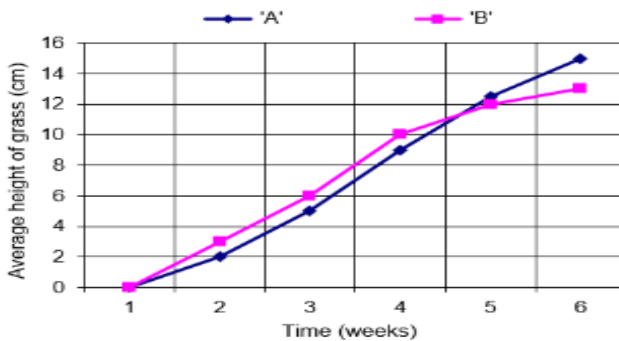
3. Maria and Hilya compared the rate of growth of two different types of grass. They planted same number of grass seeds at the same time in two identical pots with same amount of soil. They kept the pots under same conditions for six weeks and made observations at the same time each week. They recorded average height of grass in each pot. They recorded their observations in the table.

Time (weeks)	Average height of grass (cm)	
	'A'	'B'
1	0	0
2	2	3
3	5	6
4	9	8
5	12	10
6	15	13

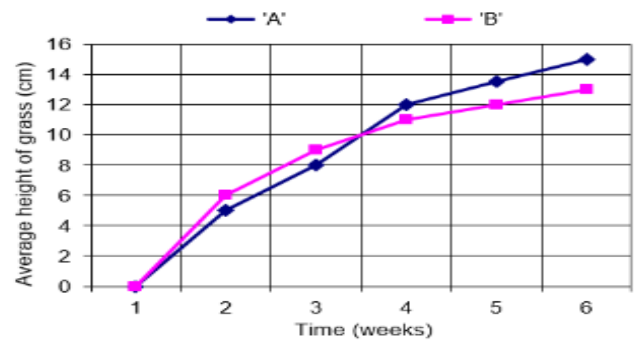
Which of the following graphs represents these results correctly?

(1)

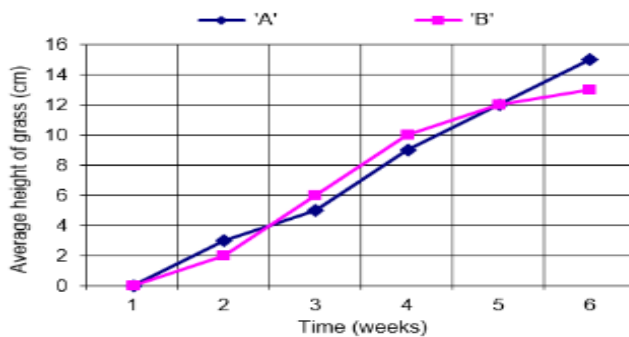
A



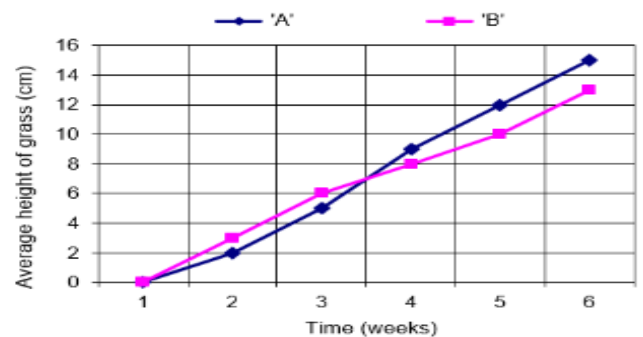
B



C



D



4. The table shows the displacement of a vertically free falling object, from the moment it was dropped from a height. Examine how far did the object move from its original position, then answer the questions. Choose from the possible answers.

Time (s)	Place from dropping (cm)
0.0	0
0.1	5
0.2	20
0.3	45
0.4	80
0.5	125
0.6	180

- 4.1 Choose an interval, when the object made the least displacement. (1)

- a) 0 – 0.1 s
- b) 0.1 – 0.2 s
- c) 0.2 – 0.3 s
- d) 0.3 – 0.4 s

- 4.2 Choose an interval, when the object reached 35 cm displacement. (1)

- a) 0.1 – 0.2 s
- b) 0.2 – 0.3 s
- c) 0.3 – 0.4 s
- d) 0.4 – 0.5 s

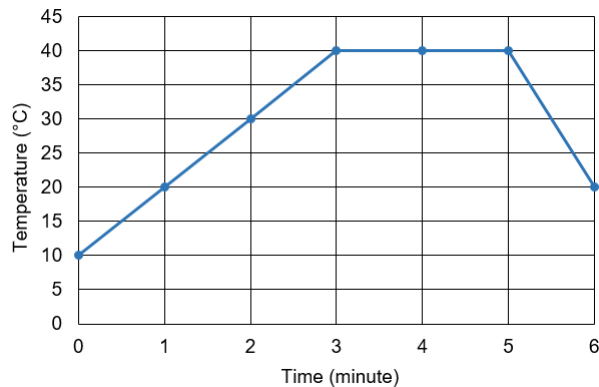
- 4.3 How does the displacement covered by the object changed every 0.1 second? (1)

- a) It increases at the same rate.
- b) It decreases at the same rate.
- c) It does not increase at the same rate.
- d) It does not decrease at the same rate.

- 4.4 How did the object move? (1)

- a) It accelerates.
- b) It gets slower.
- c) Its speed was constant.
- d) It stopped moving

5. The graph presents the change of temperature of 1 kg water with time. Use the graph to answer the questions. Choose the correct answer.



5.1 What was the temperature of the water at the end of the third minute? (1)

- a) 20°C
- b) 30°C
- c) 40°C
- d) 50°C

5.2 In the first three minutes by how many degrees, did the temperature of water change? (1)

- a) 20°C
- b) 30°C
- c) 40°C
- d) 50°C

5.3 By how many degrees did the temperature of water change in the last three minutes? (1)

- a) 20°C
- b) 30°C
- c) 40°C
- d) 50°C

5.4 How did the temperature of water changed from the starting point in 6 minutes. (1)

- a) It decreased to its half.
- b) It decreased to one quarter of the original.
- c) It increased four times higher than the original.
- d) It was doubled.

6. Students observed temperature ranges for seeds of different plants to germinate. In the table the Minimum shows the least, the Maximum shows the highest temperature at which the seeds would germinate. Optimum shows the most favourable temperature for germination.

Plants	Minimum (°C)	Optimum (°C)	Maximum (°C)
Peas	0–4.5	25–31	31–37
Alfalfa	0–4.8	31–37	37–40
Sunflower	4.8–5.5	31–37	37–44
Melon	15–18	31–37	44–50

The students draw conclusion statements from the table. Are these true or false? Choose the right answer for each.

(4)

6.1 You can plant peas the earliest in the soil.

TRUE	FALSE
------	-------

6.2 Alfalfa and melon has the same optimal temperature for germination.

TRUE	FALSE
------	-------

6.3 Sunflower germinates best between 37–44°C.

TRUE	FALSE
------	-------

6.4 Above 50°C none of these plants can germinate

TRUE	FALSE
------	-------

Total Marks for section A : 15

Section B

7. In a chemistry class students examined how temperature changes can cause the milk become sour. They poured milk of the same brand and same fat content into two identical glasses. They kept one glass at 4 °C and the other glass at 20 °C for the same period.

Analyze the factors of the experiment written below the table. Write down each factor into the appropriate box. *For example;* Write **A** with the appropriate factor. **(5)**



A	The changes of this factor was examined in the experiment	
B	This factor was changed/varied in the experiment	
C	This factor was not changed/varied during the experiment	
D	This factor was not changed/varied during the experiment	
E	This factor was not changed/varied during the experiment	

Temperature

The amount of milk

The fat content of the milk

The time of storage

The rate at which milk become sour

8. Students observed the effervescent tablets dissolving in water. They had two glasses of water at different temperature, one glass at 20°C and the other at 60°C. The volume of water in the glasses was the same. The tablets used were of the same quality and size.



Analyze the used factors below the table and write the names of each factor into the appropriate box. *For example*, write **A**, with the appropriate factor. (5)

A	This factor remained constant/unchanged during the experiment	
B	This factor remained constant/unchanged during the experiment	
C	This factor was changed/varied during the experiment	
D	This factor changed/varied during the experiment	
E	This factor changed/varied during the experiment	

The rate at which bubbles are formed

The amount of water

Temperature

The rate of dissolution

The quality of the tablets

9. In class, students studied ancient Greek ships. They conducted an experiment to decide in which direction the ship goes faster. They created models of Greek ships, and they brought these into a water filled canal.



Before the experiment:

- They tied a weight on a string to the models. The weights pulled the ships with an equal and constant force.
- They put weights into the ships to create equal load lines.

During the experiment:

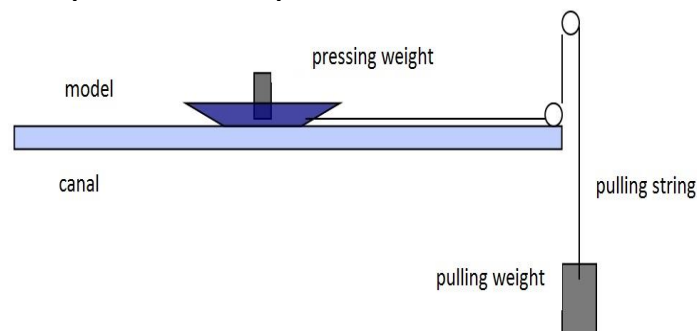
- Students moved the ships first in one direction, then changed in the other direction.
- In both cases, they measured the time it took the ship to move from one end of the canal to the other end.

What role did the following factors play in the experiment?

Write the factors (the variables) provided below, into the appropriate boxes. NB: one box needs **two factors** (variables). *For example*, write **A**, with the appropriate factor.

(3)

The experimental set up



the direction of the ship

pulling weight

pressing weight

duration of passage

Fixed variable
(its value does not change during the experiment)

A

Independent variable
(its value was changed at will during the experiment)

B

Dependent variable
(its value depended on the setting of the independent value)

C

Total Marks for section B : 13

Section C

10. Students were wondering about water uptake in plants. They made the following experiment:

They poured 100-100 ml water in graduated cylinders. They put a small amount of paraffinic oil on the water to prevent the evaporation. They put stems from the same plant into the graduated cylinders. The stems were of different size and had different numbers of leaves. They kept the cylinders at different temperature (warm and cold).



Read the questions and decide whether they can be answered with the procedure stated above or not. (4)

10.1 Does evaporation depend on the number of leaves?

YES	NO
-----	----

10.2 Does evaporation depend on the type of the plant?

YES	NO
-----	----

10.3 Does evaporation depend on the number of water?

YES	NO
-----	----

10.4 Does evaporation depend on the environment of the plant?

YES	NO
-----	----

11. Students made parachute from plastic bags and canvas and hung a toy on it. They dropped the parachute. They used several parachutes of different sizes, but the toy was always the same.



Read the questions and decide whether they can be answered with the procedure stated above or not. (4)

Does the duration of the fall depend on...

11.1 the mass of the toy?

YES	NO
-----	----

11.2 the mass of the parachute?

YES	NO
-----	----

11.3 the material of the parachute?

YES	NO
-----	----

11.4 the size of the parachute?

YES	NO
-----	----

12. Students examined the dust-pollution of their town. They put sello-tape on the leaves of acacia trees. They took the tape off, and put it carefully on a piece of glass. Then they counted the dust particles that stuck to the tape under a microscope. They examined the dust-pollution close to a busy highway and at a far distance from the highway. The leaves were always collected at two heights.



Read the questions and decide whether they can be answered with the procedure stated above or not.

(4)

12.1 Does the weather affect dust-pollution?

YES	NO
-----	----

12.2 Is there a relationship between the degree of dust-pollution and the distance from the highway?

YES	NO
-----	----

12.3 Does the degree of dust-pollution depend on the distance from the ground?

YES	NO
-----	----

12.4 Is there a difference between the dust binding of different types of trees?

YES	NO
-----	----

13. Students examined the salt being dissolved in water. They conducted the two experiments shown in the table.

Experiments	Examined	Varied	Constant
1	Quantity of salt dissolved	Temperature	Amount of water
2	Quantity of salt dissolved	Amount of water	Temperature

Read the questions below, then decide which one can be answered by the experiments above. Choose the right answer. (1)

- How does temperature affect the quantity of salt dissolved?
- How does the amount of water affect the quantity of salt dissolved?
- How do the amount of water and temperature affect the quantity of salt dissolved?
- How does the quantity of salt affect the temperature of the solution?

14. Students performed two series of experiments on factors that influence combustion conditions. They summarized the properties of the experiments in a table.

Experiments	Varied	Constant
1	Type of material burned	Temperature and the amount of oxygen
2	The amount of oxygen	Type of material burned and the temperature

Read the questions below, then decide for each questions if it can be answered by the above experiments. (4)

- 14.1 How does temperature affect onset of combustion?

YES	NO
-----	----

- 14.2 How does the quantity of the material burned affect combustion?

YES	NO
-----	----

- 14.3 How does the amount of oxygen affect the combustion?

YES	NO
-----	----

- 14.4 How does the type of material burned affect the combustion?

YES	NO
-----	----

Section D.

15. What happens if we leave a glass of water in a room for some days without covering it?



Maria said some of the water would evaporate.

John thought there would be no change at all.

Gabriel says there would be vapour on the glass.

They made the following experiment to check which statement was true. They put the same glass of water on one pan of an equal-arm balance and put some weights on the other arm pan for the arm balance to reach equilibrium state. They did not cover the glass and they kept the balance in a warm room. After two days they found that the arm of the balance (tilted) moved lower to the side of the weights.

Whose hypothesis was correct based on the experiment?

(1)

- a) Maria
- b) John
- c) Gabriel

16. Students mixed 20°C and 40°C water in a bowl. Before the experiment , they discussed what the temperature of water will be after mixing. They started the following hypotheses:

Danny thinks, that the new temperature will be the sum of the two original temperatures.

Ester thinks that the new temperature will be between the two original ones, but it will be closer to the temperature of water in larger quantity (water with more mass).

Ndina thinks that the new temperature will be the avarage of the two original ones.

After that, they made the experiment. They write the mass of water and the temperatures in a table.

Masurement	Mass of 20°C water (g)	Mass of 40°C water (g)	Temperature of the mixed water (°C)
1	50	50	30
2	80	20	24
3	20	80	38
4	10	90	36
5	90	10	22

Whose hypothesis was correct based on the experiment?

(1)

- a) Danny
- b) Ester
- c) Ndina

17. Ndeshi decided to examine whether objects with different colors absorb heat from the sun at the same rate. She poured equal amount of water into five identical glass cups. She covered the cups with the same plastic foil in different colours, black, red, blue, and white, but one glass cup was not covered. She arranged the cups so that same amount of sunlight reached each of them. After one hour, she measured the temperature of water in each cup.



Which of the following hypotheses did she test?

(1)

- a) The more sunlight heats the cups the warmer the cups become.
- b) Different kinds of materials are heated to different temperatures by sun.
- c) Different colors absorb sun light at different rate.
- d) Sunlight heats water most.

18. Ningeni wanted to investigate whether the temperature of water affect the amount of sugar dissolved in it. She poured equal amount of water into two identical cups. She then heated one cup to 20°C and the other cup one to 60°C. She put one sugar cube into each cup, and stirred until the cube dissolved. She repeated this process until cubes did not dissolve anymore. She counted how many sugar cubes were dissolved in each cup.



Which of the following hypothesis did she test?

(1)

- a) As the amount of sugar increases, so does the amount of dissolved sugar increase.
- b) As temperature of water increases, so does the amount of dissolved sugar increase.
- c) As the amount of water increases, so will the amount of dissolved sugar increases.
- d) Stirring will increase the amount of dissolved sugar.

19. Tangeni thought that the more air pressure is in a basketball, the better it will bounce. To test this hypothesis, he took two identical balls, and pumped different air pressure into them. How could Tangeni test his hypothesis? Choose the correct answer. **(1)**



- a) He should drop the ball with less air from 2 meters and the ball with more air from 1 meter, and measure how high they bounce.
- b) He should put the two balls on top of each other and drop them from 1.5 meters and measure how high they bounce.
- c) He should drop both balls from 2 meters and measure how high they bounce.
- d) He should drop the ball with less air from 1 meter and the ball with more air from 2 meters and measure how high they bounce.

20. Students observed the pressure which comes from the weight of the water. They made three holes on the plastic bottles as shown in the pictures. They covered the holes with their fingers, and filled the bottles with water. They lifted their fingers and observed how far the water coming out of the holes would reach. Before the experiment they made their hypotheses.



Would this experiment test the following hypotheses? Tick on the right answer for each hypotheses. (4)

a) If the holes are placed on a horizontal line, than the water from the holes will reach the same distance.

YES	NO
-----	----

b) If the holes are placed on a vertical line, than the water from bottom hole will reach the farthest.

YES	NO
-----	----

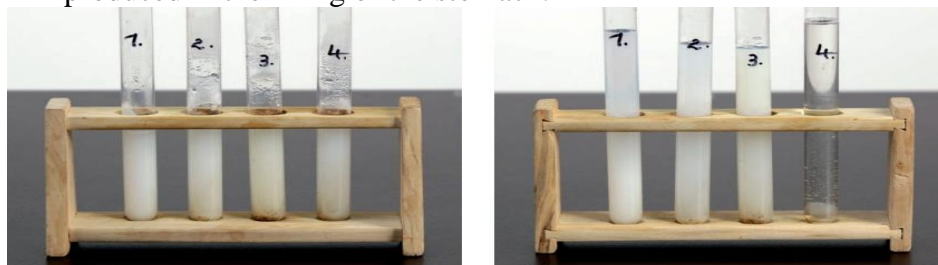
c) If the water level decreases in the bottle, than the distance covered by the water will decrease.

YES	NO
-----	----

d) If the water is warm, the water from holes will reach farther than with cold water conditons.

YES	NO
-----	----

21. In the human intestinal tract, organic nutrients (fats, oils, proteins and sugar) are decomposed by gastric juices. An experiment was performed to examine the effect of pepsin, which is produced in the lining of the stomach.



A solution of egg whites was put in four test tubes. Then materials indicated with an X in the table were added to the test tubes. 20 minutes later we found that the protein was only digested in the fourth test tube.

Experiment	Water	Weak hydrochloric acid solution	Pepsin solution
Test tube 1	X	-	-
Test tube 2	X	X	-
Test tube 3	-	-	X
Test tube 4	-	X	X

Is this experiment appropriate to verify the following statements? Tick on the answer.

(4)

Pepsin

a) can decompose protein.

YES	NO
-----	----

b) is produced in the lining of the stomach.

YES	NO
-----	----

c) is only effective in an acidic environment.

YES	NO
-----	----

d) does not decompose fats.

YES	NO
-----	----

Total Marks for section D : 13

Section E.

22. Students wanted to find out whether the rate of the evaporation of liquids depends on temperature. They dropped liquid on two watch glasses and left them uncovered on a table. One was left in a room at 20°C, the other at 25°C. The students observed how long it took for the liquids on the glasses to completely evaporate. The experiments of the student groups are shown in the table.

Which student group made the experiment appropriate to answer the question?

(1)

Student group	 watch glass 1 20°C	 watch glass 2 25°C
1	2 ml alcohol	2 ml water
2	2 ml alcohol	3 ml alcohol
3	2 ml alcohol	2 ml alcohol

23. When a rolling ball hits a stationary ball, the stationary ball will start to move. If we want to find out how its speed is influenced by the mass of the rolling ball which has hit it, we can perform several measurements.



Which two measurements should be performed to answer the question? Choose them from the table. **(1)**

	The mass of the rolling ball (kg)	The speed of the rolling ball (m/s)	The mass of the stationary ball (kg)
A	0.1	0.1	0.2
B	0.1	0.2	0.2
C	0.3	0.1	0.2
D	0.3	0.1	0.3

24. A tightly strung wire will makes sound when it is hit. Mathew and Victoria observed how the pitch depends on different factors. They made the table to show the parameters of the string in different experiments.



Experiment	Material	Thickness (mm)	Temperature (°C)	Length (cm)
A	steel	1	25	50
B	steel	2	25	40
C	steel	1	25	30
D	aluminium	1	25	40
E	aluminium	2	25	40

Which two experiments would answer the following questions? Write the letters of the experiments for each question. (3)

How does the pitch depend on...

24.1) the materials of the wire?

24.2) the length of the wire?

24.3) the thickness of the wire?

25. When we drop an object, it often makes a trace (mark) on the material where it fell. John and Alex examined what affect the depth of the trace or of the mark. They used balls of the same size. Before the experiment and listed the parameters in this table.

Experiment	Height from which the object is dropped	The material of the ball	The material on to which the ball is dropped.
A	2	iron	flour
B	2	iron	sand
C	3	plastic	gravel
D	3	wood	sand
E	3	iron	gravel
F	2	plastic	gravel

Which two experiments would answer the following questions? Write the letters of the experiments for each question. (3)

How does the depth of the trace (mark) depend on...

25.1) the material of the ball?

25.2) the material on which the ball lands?

25.3) the height from which the ball is dropped?

26. Students examined how fluids behave in capillaries (thin tubes). They poured fluid into a glass and place a tube into it. The tube is open at both ends. Students observe how high the fluid is in the tube compared to its level in the glass.



They wrote the parameters of the experiment in this table.

Experiments	Type of fluid	Material of the tube	Inner diameter of the tube (mm)
A	fruit juice	plastic	2
B	fruit juice	glass	4
C	oil	plastic	4
D	oil	glass	2
E	oil	plastic	2

Which two experiments would answer the following questions? Write the letters of the experiments for each question. (3)

How does the level of fluid in the tube depend on...

26.1) the type of fluid?

26.2) the material of the tube?

26.3) the inner diameter of the tube?

Total Marks for section E: 11

Section F.

27. There is water in a glass. Students have to measure the weight of the liquid in the glass. They have a balance scale for the measurement. They planned the measurements.



What is the appropriate order of the measurement? Write the steps into the appropriate in order!
(1)

They measure the weight
of the empty glass.

They subtract the weight of the empty
glass from the full glass.

They measure the weight of glass
full of water.

They pour the water from the glass
to the sink.

Step 1

Step 2

Step 3

Step 4

28. Rita, Johanna, Kamati and Mandume wanted to see which of them had the greatest capacity to exhale. They made the comparison with a water displacement method. They filled water into a tank. They each exhaled by blowing the air into a balloon.



Put the following steps into the **right order** from the first step to the last one. Write down the numbers only in order. (1)

1. They submerged a balloon into the tank carefully, so that it would be completely under water, but the instrument with which they kept the balloon under the surface would displace very very little water.
2. They compared and $h_2 - h_1$ differences and this gave them the answer to the question.
3. The steps were repeated for each balloon.
4. They measured the height of the water when the balloon was in it (h_2).
5. They calculated the differences between h_2 and h_1 water heights.
6. They measured the initial height of the water in the tank (h_1).

29. The hardness of rocks can be established with the following simple experimental method.



If the rock...

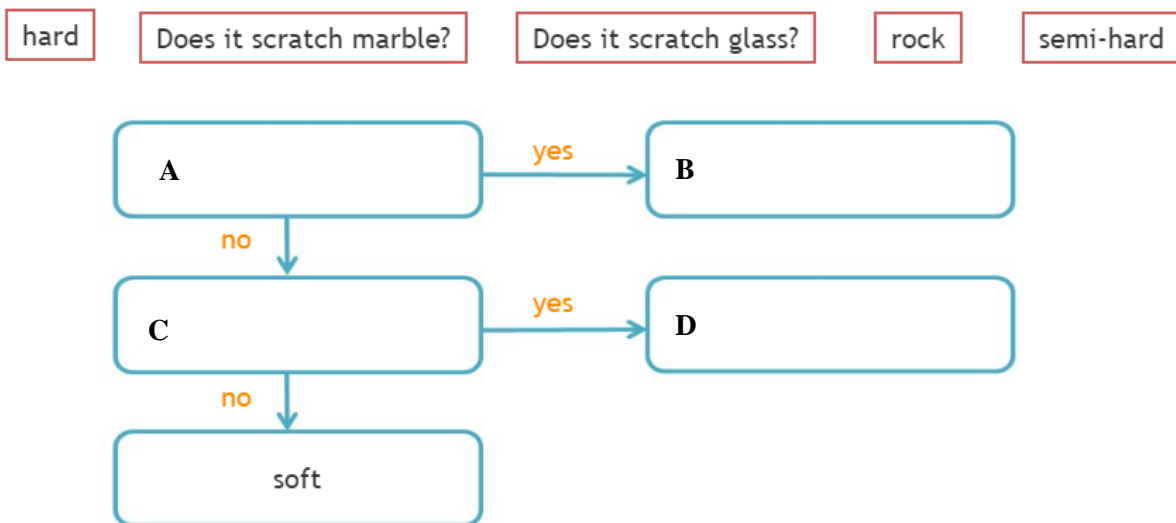
...scratches glass, it is hard;

...does not scratch glass but does scratch the marble, it is semi-hard;

...scratches neither glass nor the marble, it is soft.

Use the words below to fill in the flow chart diagram for the examination of rocks.





(4)



30. We examined how light and water affect the growing of plants. We assume that plants need light and water both to stay alive. Under what conditions should we keep the plants to prove this hypothesis? Create the experiment conditions.

Write the names of the environmental factors below for each experiments. Remember each experiment need **two** conditions. (4)



Experiment 1	Experiment 2	Experiment 3	Experiment 4
			
A	B	C	D

Total Marks for section F : 10

Section G.

31. Danny poured water into a test tube and added a teaspoon of starch to it. He shook the testtube and add some dops iodine solution to it. The mixture turned blue.

He then cut a potato into half and put some drops of iodine solution on the potato. A blue spot appeared on the potato.



What does this experiment verify? Choose the right answer. (1)

- a) Iodine is soluble in starch.
- b) Potato contains iodine.
- c) Potato contains starch.
- d) Potato contains water.

32. There are 3 conditions needed to be satisfied for combustion to occur: i) there should be combustible material available, ii) the material should reach the combustion temperature, iii) there should be enough oxygen. We made the following experiments to verify those conditions.



Which experiments were verified by which conditions?

Write each letter of the experiment into the appropriate box. (4)

- a) We lit a candle then we put it into a heat resistant glass container with a glass lid. The candle slowly went off.
- b) We put a piece of limestone into the flame of a Bunsen burner. The limestone did not burn.
- c) We put a piece of paper out to the sunlight. The paper did not catch fire.
- d) We put a burning piece of paper on an iron plate. The flame went off when there was only ash on the plate.

32.1	Combustible material	
32.2	Combustible material	
32.3	Combustion temperature	
32.4	Oxygen	

33. Students compared the different states of matter of water. The ice tray was filled with water to the top (brim). The tray was put in a freezer. The ice cubes formed were higher than the top (brim) of the tray. They made the following table from their result.



Experiment	Activity	Result
1	They filled completely the ice tray, then put it in the freezer.	Ice cubes overflowing out from the ice-tray.
2	Ice cubes are put into water.	Ice cubes floated on the surface of water and their size was continually decreasing.

Could we make the following conclusions based on these experiments?

Choose the right answer for each conclusion.

(4)

a) The freezing point of water and the melting point of ice are the same.

YES	NO
-----	----

b) If water is frozen, its volume will change.

YES	NO
-----	----

c) Ice cubes melt at room temperature.

YES	NO
-----	----

d) Ice has smaller density than water.

YES	NO
-----	----

34. Students observed the solubility of materials. They arranged their observations in a table.

Experiment	Activity	Observation
1	We pour a small amount of oil into the water in a test tube then we shake it.	The oil accumulates on top of the water.
2	We pour a small amount of petrol into the alcohol in a test tube then we shake it.	We see a colourless liquid in the test tube, and we cannot differentiate the two materials.

Could they make the following conclusions based on these experiments?

Choose the right answer for each conclusion.

(4)

a) The density of water is higher than the density of oil.

YES	NO
-----	----

b) The density of alcohol is lower than the density of petrol.

YES	NO
-----	----

c) Oil does not dissolve in water.

YES	NO
-----	----

d) Alcohol dissolves in water.

YES	NO
-----	----

- 35.** Moses and Nelago were told to compare the density of different materials. They already knew that solid objects float on the surface of liquids if their density is smaller than that of the liquid; they sink if their density is bigger than that of the liquid; and they float in the liquid if the densities of the solid object and the liquid are the same.

Moses and Nelago put different solid objects into different liquids. This is what they observed:

The wooden ball floated on water.

The wooden ball floated on oil.

The aluminium ball sank in water.

The aluminium ball sank in oil.

Can they draw the following conclusions from their observations?

(4)

- a) The density of wood is smaller than the density of oil.

YES	NO
-----	----

- b) The density of oil is smaller than the density of water.

YES	NO
-----	----

- c) The density of wood is smaller than the density of aluminium.

YES	NO
-----	----

- d) The density of water is smaller than the density of aluminium.

YES	NO
-----	----

36. The students created a battery out of a fruit, a vegetable and two pieces of metal. They measured the voltage created by the battery. They tested several conditions and observed the following:

Jonas connected a magnesium and a copper rod to a lemon and measured 1.6 V.

Gloria connected a zinc and a copper rod to a lemon and measured 0.9 V.

Olivier connected two iron nails to a potato and measured 0 V.

Kate connected a zinc and a copper rod to a potato and measured 1.1 V.

Could we draw the following conclusions based on their measurements? Choose the right answer for each conclusion. **(4)**

a) Voltage depends on the material of metals.

YES	NO
-----	----

b) Voltage does not depend on the material of metals.

YES	NO
-----	----

c) Voltage does not depend on the size of metals.

YES	NO
-----	----

d) Voltage depends on the fruit or the vegetable.

YES	NO
-----	----

Total Marks for section G : 21

Total Marks for the paper : 100
--

Please, tell us about your parents' education. What is the highest schooling level they completed? If you live with a step-parent, give information about his or her education. Choose your answer from the list below.



BG3. Mother highest level of education:

- | | |
|----------------------------------|--------------------|
| a) Did not go to school. | e) Bachelor degree |
| b) Did not finish primary school | f) Master degree |
| c) Primary school | g) PhD degree |
| d) Secondary education | h) I do not know |

BG4. Father highest level of education:

- a) Did not go to school.
- b) Did not finish primary school
- c) Primary school
- d) Secondary education
- e) Bachelor degree
- f) Master degree
- g) PhD degree
- h) I do not know

You have reached the END of the test. Thank you so much!!

Appendix D: Lawson Classroom Test of Scientific reasoning (8th, 10th & 12th grades)

CLASSROOM TEST OF SCIENTIFIC REASONING

Directions to Students:

This is a test of your ability to apply aspects of scientific analyze a situation to make a prediction or solve a problem. Make a dark mark on the answer sheet for the best answer for each item. If you do not fully understand what is being asked in an item, please ask the test administrator for clarification.

DO NOT OPEN THIS BOOKLET UNTIL YOU ARE TOLD TO DO SO

Revised Edition: August 2000 by Anton E. Lawson, Arizona State University. Based on: Lawson, A.E. 1978. Development and validation of the classroom test of formal reasoning. *Journal of Research in Science Teaching*, 15(1): 11-24.

1. Suppose you are given two clay balls of equal size and shape. The two clay balls also weigh the same. One ball is flattened into a pancake-shaped piece. Which of these statements is correct?

- The pancake-shaped piece weighs more than the ball
- The two pieces still weigh the same
- The ball weighs more than the pancake-shaped piece

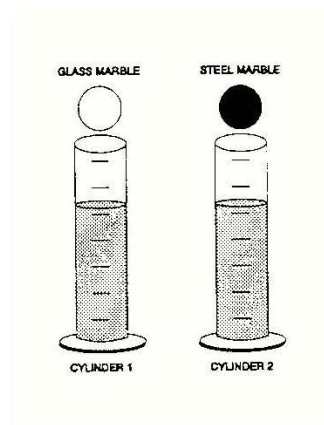
2. *because*

- the flattened piece covers a larger area.
- the ball pushes down more on one spot.
- when something is flattened it loses weight.
- clay has not been added or taken away.
- when something is flattened it gains weight.

3. To the right are drawings of two cylinders filled to the same level with water. The cylinders are identical in size and shape.

Also shown at the right are two marbles, one glass and one steel. The marbles are the same size but the steel one is much heavier than the glass one.

When the glass marble is put into Cylinder 1 it sinks to the bottom and the water level rises to the 6th mark. *If we put the steel marble into Cylinder 2, the water will rise*



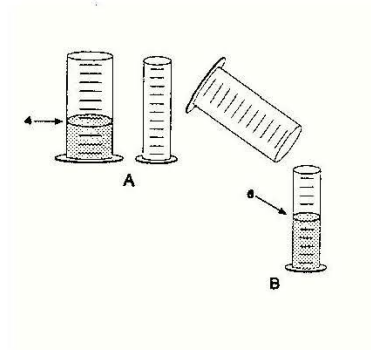
- to the same level as it did in Cylinder 1
- to a higher level than it did in Cylinder 1
- to a lower level than it did in Cylinder 1

4. *because*

- the steel marble will sink faster.
- the marbles are made of different materials.
- the steel marble is heavier than the glass marble.
- the glass marble creates less pressure.
- the marbles are the same size.

5. To the right are drawings of a wide and a narrow cylinder. The cylinders have equally spaced marks on them. Water is poured into the wide cylinder up to the 4th mark (see A). This water rises to the 6th mark when poured into the narrow cylinder (see B).

Both cylinders are emptied (not shown) and water is poured into the wide cylinder up to the 6th mark. *How high would this water rise if it were poured into the empty narrow cylinder?*



- to about 8
- to about 9
- to about 10
- to about 12
- none of these answers is correct

6. *because*

- the answer can not be determined with the information given.
- it went up 2 more before, so it will go up 2 more again.
- it goes up 3 in the narrow for every 2 in the wide.
- the second cylinder is narrower.
- one must actually pour the water and observe to find out.

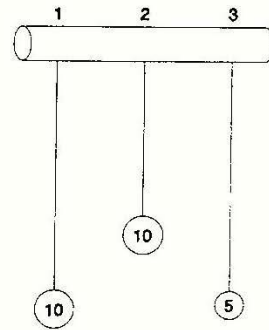
7. Water is now poured into the narrow cylinder (described in Item 5 above) up to the 11th mark. *How high would this water rise if it were poured into the empty wide cylinder?*

- to about $7 \frac{1}{2}$
- to about 9
- to about 8
- to about $7 \frac{1}{3}$
- none of these answers is correct

8. *because*

- the ratios must stay the same.
- one must actually pour the water and observe to find out.
- the answer can not be determined with the information given.
- it was 2 less before so it will be 2 less again.
- you subtract 2 from the wide for every 3 from the narrow.

9. At the right are drawings of three strings hanging from a bar. The three strings have metal weights attached to their ends. String 1 and String 3 are the same length. String 2 is shorter. A 10 unit weight is attached to the end of String 1. A 10 unit weight is also attached to the end of String 2. A 5 unit weight is attached to the end of String 3. The strings (and attached weights) can be swung back and forth and the time it takes to make a swing can be timed.



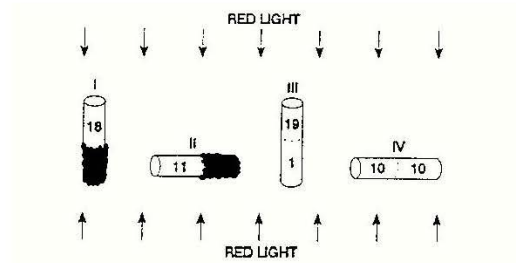
Suppose you want to find out whether the length of the string has an effect on the time it takes to swing back and forth. *Which strings would you use to find out?*

- only one string
- all three strings
- 2 and 3
- 1 and 3
- 1 and 2

10. *because*

- you must use the longest strings.
- you must compare strings with both light and heavy weights.
- only the lengths differ.
- to make all possible comparisons.
- the weights differ.

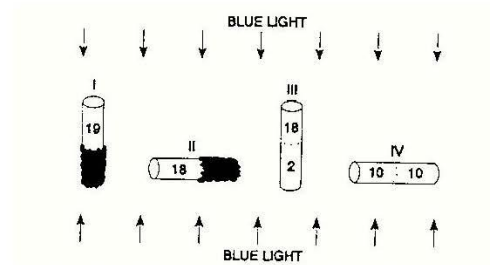
11. Twenty fruit flies are placed in each of four glass tubes. The tubes are sealed. Tubes I and II are partially covered with black paper; Tubes III and IV are not covered. The tubes are placed as shown. Then they are exposed to red light for five minutes. The number of flies in the uncovered part of each tube is shown in the drawing.



This experiment shows that flies respond to (respond means move to or away from):

- red light but not gravity
 - gravity but not red light
 - both red light and gravity
 - neither red light nor gravity
12. *because*
- most flies are in the upper end of Tube III but spread about evenly in Tube II.
 - most flies did not go to the bottom of Tubes I and III.
 - the flies need light to see and must fly against gravity.
 - the majority of flies are in the upper ends and in the lighted ends of the tubes.
 - some flies are in both ends of each tube.

13. In a second experiment, a different kind of fly and blue light was used. The results are shown in the drawing.



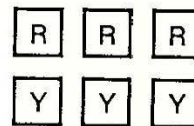
These data show that these flies respond to (respond means move to or away from):

- blue light but not gravity
- gravity but not blue light
- both blue light and gravity
- neither blue light nor gravity

14. *because*

- some flies are in both ends of each tube.
- the flies need light to see and must fly against gravity.
- the flies are spread about evenly in Tube IV and in the upper end of Tube III.
- most flies are in the lighted end of Tube II but do not go down in Tubes I and III.
- most flies are in the upper end of Tube I and the lighted end of Tube II.

15. Six square pieces of wood are put into a cloth bag and mixed about. The six pieces are identical in size and shape, however, three pieces are red and three are yellow. Suppose someone reaches into the bag (without looking) and pulls out one piece. *What are the chances that the piece is red?*

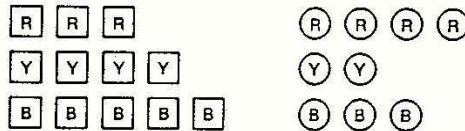


- 1 chance out of 6
- 1 chance out of 3
- 1 chance out of 2
- 1 chance out of 1
- cannot be determined

16. *because*

- a. 3 out of 6 pieces are red.
- b. there is no way to tell which piece will be picked.
- c. only 1 piece of the 6 in the bag is picked.
- d. all 6 pieces are identical in size and shape.
- e. only 1 red piece can be picked out of the 3 red pieces.

17. Three red square pieces of wood, four yellow square pieces, and five blue square pieces are put into a cloth bag. Four red round pieces, two yellow round pieces, and three blue round pieces are also put into the bag. All the pieces are then mixed about. Suppose someone reaches into the bag (without looking and without feeling for a particular shape piece) and pulls out one piece.



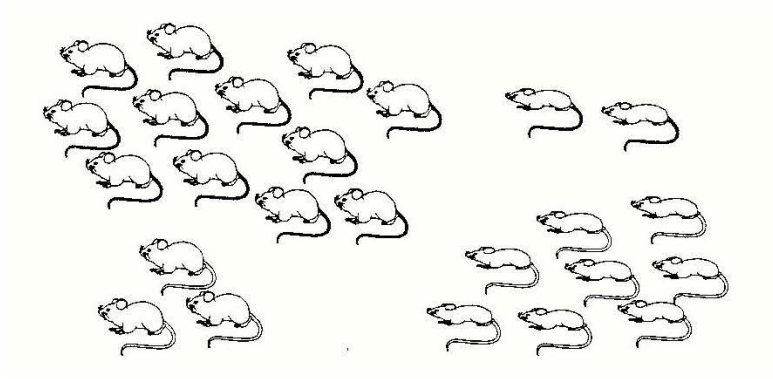
What are the chances that the piece is a red round or blue round piece?

- a. cannot be determined
- b. 1 chance out of 3
- c. 1 chance out of 21
- d. 15 chances out of 21
- e. 1 chance out of 2

18. *because*

- a. 1 of the 2 shapes is round.
- b. 15 of the 21 pieces are red or blue.
- c. there is no way to tell which piece will be picked.
- d. only 1 of the 21 pieces is picked out of the bag.
- e. 1 of every 3 pieces is a red or blue round piece.

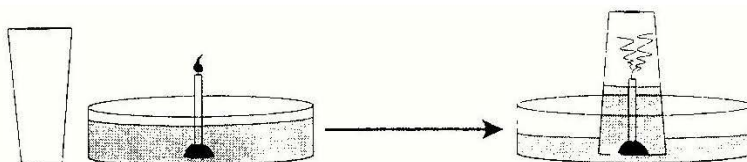
19. Farmer Brown was observing the mice that live in his field. He discovered that all of them were either fat or thin. Also, all of them had either black tails or white tails. This made him wonder if there might be a link between the size of the mice and the color of their tails. So he captured all of the mice in one part of his field and observed them. Below are the mice that he captured.



Do you think there is a link between the size of the mice and the color of their tails?

- a. appears to be a link
 - b. appears not to be a link
 - c. cannot make a reasonable guess
20. *because*
- a. there are some of each kind of mouse.
 - b. there may be a genetic link between mouse size and tail color.
 - c. there were not enough mice captured.
 - d. most of the fat mice have black tails while most of the thin mice have white tails.
 - e. as the mice grew fatter, their tails became darker.

21. The figure below at the left shows a drinking glass and a burning birthday candle stuck in a small piece of clay standing in a pan of water. When the glass is turned upside down, put over the candle, and placed in the water, the candle quickly goes out and water rushes up into the glass (as shown at the right).



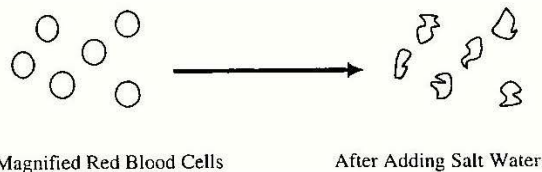
This observation raises an interesting question: Why does the water rush up into the glass?

Here is a possible explanation. The flame converts oxygen into carbon dioxide. Because oxygen does not dissolve rapidly into water but carbon dioxide does, the newly formed carbon dioxide dissolves rapidly into the water, lowering the air pressure inside the glass.

Suppose you have the materials mentioned above plus some matches and some dry ice (dry ice is frozen carbon dioxide). *Using some or all of the materials, how could you test this possible explanation?*

- Saturate the water with carbon dioxide and redo the experiment noting the amount of water rise.
 - The water rises because oxygen is consumed, so redo the experiment in exactly the same way to show water rise due to oxygen loss.
 - Conduct a controlled experiment varying only the number of candles to see if that makes a difference.
 - Suction is responsible for the water rise, so put a balloon over the top of an open-ended cylinder and place the cylinder over the burning candle.
 - Redo the experiment, but make sure it is controlled by holding all independent variables constant; then measure the amount of water rise.
22. What result of your test (mentioned in #21 above) would show that your explanation is probably wrong?
- The water rises the same as it did before.
 - The water rises less than it did before.
 - The balloon expands out.
 - The balloon is sucked in.

23. A student put a drop of blood on a microscope slide and then looked at the blood under a microscope. As you can see in the diagram below, the magnified red blood cells look like little round balls. After adding a few drops of salt water to the drop of blood, the student noticed that the cells appeared to become smaller.



This observation raises an interesting question: Why do the red blood cells appear smaller?

Here are two possible explanations: I. Salt ions (Na^+ and Cl^-) push on the cell membranes and make the cells appear smaller. II. Water molecules are attracted to the salt ions so the water molecules move out of the cells and leave the cells smaller.

To test these explanations, the student used some salt water, a very accurate weighing device, and some water-filled plastic bags, and assumed the plastic behaves just like red-blood-cell membranes. The experiment involved carefully weighing a water-filled bag, placing it in a salt solution for ten minutes and then reweighing the bag.

What result of the experiment would best show that explanation I is probably wrong?

- a. the bag loses weight
 - b. the bag weighs the same
 - c. the bag appears smaller
24. *What result of the experiment would best show that explanation II is probably wrong?*
- a. the bag loses weight
 - b. the bag weighs the same
 - c. the bag appears smaller

Appendix E: English Reading comprehension test (8th grade)

Name.....Grade...

Task 1: Find the pairs of these definitions from the list (A - K). There is an extra word you do not need. See the example.

- A** A person you visit if you have a headache or a sore throat.
- B** A person who helps you with tickets and information.
- C** This is what you do if you want to cross the ocean or travel a long way fast.
- D** A person who takes criminals to prison.
- E** This is a vehicle you take if you want to get from one town to the other.
- F** A place where you visit someone sick or ill.
- G** A person who delivers letters.
- H** You do this when you are tired and you want to relax.
- I** You can do this in a park, street or forest.
- J** You can do this if you want to see far and there is a tree near you.
- K** A place where you go to buy your ticket if you want to travel somewhere.

- | | | | |
|-------------------|--------------------|------------------|------------------------|
| 1 <i>doctor</i> | 4 <i>watch</i> | 7 <i>station</i> | 10 <i>travel agent</i> |
| 2 <i>hospital</i> | 5 <i>policeman</i> | 8 <i>sleep</i> | 11 <i>walk</i> |
| 3 <i>fly</i> | 6 <i>postman</i> | 9 <i>climb</i> | 12 <i>train</i> |

Put your answers in here:

A	B	C	D	E	F	G	H	I	J	K
1										

Task 2: Here are some notices. What do they mean? Choose from the answers 1-12. There is one extra you do not need. See the example.

- A** DON'T TOUCH WINDOW
- B** *Opening hours: 9 am - 5 pm*
- C** *TWO HOUR PHOTO SERVICE IN HERE*
- D** **Please keep it open**
- E** *No eating or drinking in the classroom*
- F** **Entrance**
- G** *TODAY: 2 for the price for 1*
- H** **PLEASE USE OTHER DOOR**
- I** **special student price! Ask at the desk**
- J** *Use before 01 February 2002*

- 1** if you want a cheap ticket ask here
- 2** make sure you are not hungry or thirsty in here
- 3** this is the place where you can go in
- 4** this is the place where you cannot go in
- 5** you can visit this place for eight hours
- 6** after this date you must not eat this food
- 7** do not go near this object!
- 8** you can buy books here
- 9** do not shut this
- 10** you can get your pictures fast in here
- 11** you can get everything for half price for one day only
- 12** this is the library desk

Put your answers in here:

A	B	C	D	E	F	G	H	I	J
7									

Task 3: Read this short interview from a British youth magazine. Match the questions (A – J) with the answers (1 – 11). There is an extra answer you do not need. See the example.

What is it like to be 12 in another country? We put the question to someone in the USA

- A** What is your name?
- B** What do you do?
- C** Where do you live?
- D** Describe a typical day.
- E** What do you do in your spare time?
- F** What is your favourite TV programme?
- G** What is your favourite magazine?
- H** Your favourite food?
- I** Anything you hate?
- J** Who are your best friends?

- 1** Up by 7am, go to school till 2.30pm. Some days I'll go out with friends till 5.30pm, some days I'll go to the movie. Then I'll be with my family for dinner, do my school work and go to bed by 10pm.
- 2** I don't read any. I prefer books like Harry Potter and the Philosopher's Stone.
- 3** Laura Hilius.
- 4** Anything sweet, I love cakes, and ice-creams.
- 5** Hang out with friends, play computer games, talk on the phone, listen to music....
- 6** I don't really like maths and science.
- 7** Lake Forest, CA, with my mom, dad, sisters, and Sneeze, our dog.
- 8** My sisters: Cathy, 13, and Jane, 15.
- 9** Fish, smells awful...
- 10** Nature films on animals, and science fiction.
- 11** I go to Pinewood School.

Put your answers in here:

A	B	C	D	E	F	G	H	I	J
3									

Task 4: Read the following quiz questions (A - I) and find the right answer (1-10). There is one extra answer you do not need. See the example.

- A** Why is it hot or cold in some places?
- B** Which is the biggest bird of prey?
- C** Why do some flowers smell sweet?
- D** Which cat runs the fastest?
- E** Which is the biggest snake?
- F** Why are many tundra flowers red or yellow?
- G** Which is the most dangerous snake?
- H** How different are our pet cats from wild cats?
- I** Why are rainforests cut down?

1 They are destroyed so the land can be used for crops, or for grazing. Tropical soils are fertile and cocoa and sugar cane can be grown after the trees have been cut down.

2 There are relatively few bees this far north, and their main pollinators are flies. Flies cannot distinguish colours like those that bees can, so the flowers do not need to be so colourful.

3 When the sun is low in the sky, its rays hit the earth at an angle. When that happens, we do not feel the sun's warmth.

4 It is a way of attracting bees, flies, butterflies and other insects.

5 The Andean condor measures up to 110 cm long and weighs up to 12 kg. Its wingspan is over 3 metres.

6 The reticulated python lives in parts of Southeast Asia. It grows to an amazing 10 metres.

7 They dig beneath the snow with their hooves and antlers to seek out tender lichens, mosses and grasses.

8 The saw-scaled carpet viper is extremely aggressive and its poison can kill humans. It lives in Africa and Asia.

9 They have exactly the same body structure and skeleton. Both rely heavily on smell for information about the world. All are meat-eaters and cannot live on a diet of plant food.

10 The cheetah is one of the speediest of all animals over short distances. It has been timed running at 105 kph over 100 metres.

(Source: Giant Book of Questions and Answers. 1999. Bath: Parragon.)

Put your answers in here:

A	B	C	D	E	F	G	H	I
3								

Task 5: Read the following advertisements and find the missing bit from the list for each gap. Put the numbers in the box next to the missing bit. See the example.

<p>Neel Akash Tandoori Restaurant</p> <p>Situated in one of London's most popular streets, serving fine North Indian & Bangladeshi cuisine in elegant & cosy surroundings.</p> <p>93 Charlotte st, London W1 Tel.: +0207637 0050 www.neel-akash.com</p> <p>1</p>	<p>THE SOCIETY OF LONDON THEATRE</p> <p>The only official half price ticket booth</p> <p>2</p> <p>Monday-Saturday 12.00-6.30PM CALL +44 161 248 8282</p>	<p>3</p> <p>71 SHEPERDS BUSH ROAD HAMMERSMITH, LONDON W6</p> <p>Book now on Tel.: 020 1253 789240 Fax.: 020 1253 789240</p> <p>Single £45 Double/Twin £65 Breakfast included</p>	<p>sun, sea and ... sand</p> <p>Looking for the holiday of a lifetime with parties, and sun-drenched beaches?</p> <p>Brochure order: 07000 007 007 007</p> <p>look no further than Club Med www.clubmed.hu</p> <p>4</p>
<p>5</p> <p>FREEPOST CY1061 WEST SUSSEX NO STAMP NEEDED</p> <p>For £21 for 12 issues – saving £4 off the full price and receive a free gift</p> <p>Please allow up to 6 weeks to delivery of your first issue, and up to 28 days for delivery of your gift</p> <p>RING 0122444474887</p>	<p>DO YOU WANT A BETTER JOB? With an ICS course it's easy to study at home for the skills and qualifications you need to impress employers.</p> <p>For more information call today</p> <p>free phone: 0500 585 6565</p> <p>6</p> <p>Choose from over 130 courses in your own time.</p> <p>For a new career</p>	<p>Websites universities & colleges agony aunt horoscopes music & charts cinema & films sports results & news personal ads gossip stories chat fun</p> <p>The best and only site for teenagers!!!!</p> <p>7</p> <p>Check it out today!</p>	<p>LEARNING FRENCH & LOST FOR WORDS?</p> <p>wordprof French will improve your vocabulary and spelling</p> <p>EVEN TELLS YOU HOW MANY WORDS TO LEARN EACH DAY!</p> <p>Tel.: 0171 494 2929 Fax.: 0171 494 2922</p> <p>8</p>

<p>Child Care America The legal way to live in the U.S.A. For a free prospectus contact: GW, 22 Upbrook Mews, London W2 3HG <u>GW@childcare.com</u> Learn about America</p> <p>9</p>	<p>Music Management Company seeks 10 With strong original material for professional development Send info and demo to Slapback Management, 27 Sherbourne Drive, Maidenhead SL6 3EO</p>
--	--

Put your answers in here:

A	BEST VALUE FOR TODAY'S EVENING PERFORMANCES	2	F	DALMACIA HOTEL	
B	Learn more. Earn more.		G	9000 words & 350 topics for PC	
C	www.thepad-uk.com		H	Open for Lunch and Dinner	
D	YES, I WANT TO SUBSCRIBE TO <i>TIME OUT</i>		I	solo artists & bands	
E	Experience different cultures		J	or contact your travel agent	