Journal of Studies in Education
ISSN 2162-6952
2019, Vol. 9, No. 4

Online Assessment of Scientific Reasoning Skills and Motivation to Learn Science among Grade 5 and 7 Students in Northern Namibia

Linus Kambeyo

Faculty of Education, Department of Mathematics, Science and Sport Education

Katima Mulilo Campus, University of Namibia

E-mail: lkambeyo@unam.na

Benő Csapó

Doctoral School of Education, University of Szeged, Szeged, Hungary

Address: H-6722 Szeged, Petofi Sgt. 30-34

E-mail: csapo@edpsy.u-szeged.hu

Received: Oct. 9, 2019 Accepted: Nov. 15, 2019 Published: November 17, 2019

doi:10.5296/jse.v9i4.15598 URL: https://doi.org/10.5296/jse.v9i4.15598

Abstract

This study aimed to assess students' scientific reasoning skills using an online assessment method and explore the relationship between their reasoning skills and motivation to learn science. Research participants were 270 Grade 5 and 346 Grade 7 students in the Oshana region of Namibia. The online reasoning skills test consisted of 36 items with 16 tasks assessing conservation, proportional, correlational and probabilistic reasoning, and logical operations in a science context. The five point Likert scale Science Motivation Questionnaire II consisted of 25 items with five subscales. Tasks were developed within and delivered by the eDia platform via the Internet. The reliability of the reasoning skills test was acceptable (Cronbach's alpha=.74), and it was very good for the Science Motivation Questionnaire (Cronbach's alpha=.91). The reasoning skill tasks were moderately difficult for the students: M=40.56%; SD=13.47%. One-parameter Rasch analyses showed that there were few items to differentiate students at the low skill levels. Task analysis showed major obstacles in students' reasoning skills for science learning. Students reported that they were moderately motivated to learn science. A weak correlation was found between the tested scientific reasoning skills and motivation to learn science ($r=.21$, $p<.01$). The study suggests that the basic ICT infrastructures in Namibian schools should be improved to exploit the advantages of online assessment.

Keywords: Assessment, online assessment, scientific reasoning, motivation to learn science, Namibia

1. Introduction

The availability and the usage of technology in recent years is not just limited to entertainment, project designs, analysis, but it is also included in education like e-learning or edutainment. These new methods take the advantage from the traditional ones which was based on "drill and practice only" (Brom et al., 2009). The advantages of digital-based learning in enhancing students' involvement and creating a combination between learning and entertainment attracted the attention of educators and researchers from across the globe (Prensky, 2007), by that, computerized testing becomes an interesting area of educational evaluation to be developed rapidly (Csapó et al., 2012).

Furthermore, in recent years, a large number of studies have highlighted the importance and benefits of technology-based assessment (TBA). A broad range of instruments, including observation protocols, tests and item banks, are available which can be used to assess different aspects of general cognitive development as well as specific skills, such as scientific reasoning skills and general thinking skills, which learners are expected to master at school. Large-scale online tests developed for specific purposes have already been successfully applied in the broadest age range, from the very beginning of primary school up to college level (Csapó & Molnár, 2017).

Technology provides great potential to assess thinking skills due to its innovative characteristics, such as novel item design, including the utilization of multimedia and simulation as well as a broad range of response capture not possible on paper (e.g. drag-and-drop, interacting with dynamic systems etc.) and automated data processing (Csapó & Molnár, 2017). Nevertheless, TBA has its own constraints, which include the lack of government-provided basic technology infrastructure, and lack of consensus whether to move entirely to online assessment among the stakeholders involved in public education.

Therefore, in this study, students' scientific reasoning skills and their motivation of learning science in the Oshana region of Namibia were assessed using an online assessment instruments. The correlation between students' motivation to learn science and their scientific reasoning was also explored. It is argued that students with a high motivation to learning science are likely to aim for good science grades and to take up science-related careers (Bryan, Glyn, & Kittleson, 2011). A study by Shaakumeni and Csapó (2018) indicates that Namibian students are motivated to learn science even though their scores in *self-efficacy* and *active learning* strategies are low. It is hoped that the results from this study will inform the science education system in Namibia of the level of scientific reasoning skills among primary students. In addition, these findings may inform practitioners of new methods of assessment and the need to further support a culture of critical thinking in students.

2. Background information

Science, technology, engineering and mathematics (STEM) is a curriculum based on the idea of educating students in four specific disciplines; science, technology, engineering and mathematics in an interdisciplinary and applied approach (Csapó et al., 2012). Rather than teach the four disciplines as separate and discrete subjects, STEM integrates them into a cohesive learning paradigm based on real-world applications. This study links three rapidly developing areas of educational research and places them in the context of the development of the Namibian education system. First, improving science education to attract young people to the Science, Technology, Engineering, and Mathematics (STEM) professions is a goal in many countries as the supply of young professionals graduating in these areas fails to meet the demands of modern economies (Csapó & Molnár, 2017). One of the main emphases in this area is the quality of science education, especially its contribution to the improvement of students' higher order thinking skills. Second, educational assessment has received growing attention both in research and practice. If certain psychological constructs are made measurable, this will pave way for conducting precise training experiments, while feedback provided by the assessment may orient practice. Third, testing has become technology-based, making even more constructs measurable, while reducing the cost and timeframe of the assessments. "As the world becomes more technologically dependent and society becomes increasingly global, an understanding of reasoning ability becomes increasingly important, particularly for developing countries that need to focus on building their STEM work force" (Ibrahim, Ding, Mollohan & Stammen, 2016, p.93)

2.1 Scientific Reasoning

There are several interpretations and definitions of scientific reasoning skills that we may draw from for establishing an assessment framework., Scientific reasoning is defined as 'formal reasoning' (Piaget & Gabain, 1965) or 'critical thinking', which represents the ability to systematically explore a problem, formulate and test hypotheses, control and manipulate variables, and evaluate experimental outcomes (Bao, Cai, Koening & Fang, 2009; Zimmerman, 2007;). Morris, Croker, Masnick, and Zimmerman (2015), stated that scientific reasoning also encompasses the reasoning and problem-solving skills involved in generating, testing and revising hypotheses or theories, and in the case of fully developed skills, reflecting on the process of knowledge acquisition and knowledge change that results from such inquiry based activities. Scientific reasoning is represented in educational standards all over the world, and it is undoubtedly one of the most relevant twenty-first-century skills (Stiller, Hartmann, Matheius, Straube, Tiemann, Nordmeier, Krüger & Upmeier zu Belzen, 2016). In the same vein, it is argued that, conceptual understanding of scientific methods and the acquisition of scientific reasoning competencies during teaching and learning of science are, therefore, highly relevant (Stiller et al., 2016). Several authors have articulated that scientific reasoning as a major component in technology-enriched societies (Fischer et al. 2014; Schwartz, Lederman, & Crawford, 2004). Furthermore, tests in science reasoning skills is imperative as teachers will be able to evaluate, reflect 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 order to develop the reasoning skills in science effectively, the need to explore and ascertain students' scientific reasoning abilities by means of assessment is of paramount importance as well as the need to explore students' motivation to learn science as the background variables. Many studies have been conducted on assessing students' scientific knowledge, and general reasoning skills mostly in developed countries (Bao et al., 2009;

Lawson, 2000; Ibrahim et al., 2016; Zimmerman, 2007).

2.2 Reasoning skills and science teaching in Namibia

Teaching science in Namibia at the primary phase is competency based. The assessment is also competency based, and general thinking and reasoning skills are not explicitly taught and encouraged. When teaching, teachers tend to focus on the end-of-term results or end-of-year results. From the National Broad Curriculum of Education (NCBE) (NIED, 2009) it can be observed that the reasoning skills are automatically embedded in the competencies to be taught. One might then ask if teachers try to incorporate the reasoning skills in their teaching or they just focused on the competencies. Several empirical studies have demonstrated that learning sciences does not result necessarily in better scientific reasoning (Adey & Csapo, 2012). For example, (Bao et al., 2009) compared Chinese and American university students' physics knowledge and scientific reasoning. They have found that although Chinese students performed much better on the science knowledge test (attributable to their more demanding high school science studies), their performance on the science reasoning test was similar to that of their American peers (Bao et al., 2009).

Therefore, the training in how to teach both science content and reasoning skills is of paramount importance, (NCBE, 2010), the approach to teaching, learning and assessment is embedded on pedagogical theories, i.e. "preparation for a knowledge-based society requires a learner-centred approach to teaching and learning. The point of departure is always what the learners already know and can do, then acquiring new knowledge through ways of working which are relevant and meaningful for them, and learning how to apply their knowledge creatively and innovatively" (p. 4). However, the national curriculum stressed that, knowledge is not learnt for its own sake, but must always lead to new understanding and new skills and the creation of new knowledge. At each step of the way, students must show how competent they are in what they understand and can do (NCBE, 2010). Psychological processes of thinking and reasoning skills play significant role in understanding science and application of knowledge in new contexts (Molnar, 2011). Its modifiability has been demonstrated in a number of training experiments (Hamers, de Koning, & Sijtsma, 1998; Molnár, 2011).

The end of school year examination does not fully address the assessment reasoning skills at any level of schooling despite calls to prioritize science, technology and innovation in the country *(Namibia vision 2030* NCBE, 2010). In primary phase, the only scientific and reliable assessment that attempts to measure students' reasoning and thinking skills are the Standardized Achievement Tests (SATs) that were introduced in 2009 (Iipinge & Likando, 2012). However, these tests too fell short of assessing scientific and general reasoning skills in the broader sense of the concept (Wenning, 2006) as they only assess students' achievement of disciplinary science upon completion of Grade 7 science curriculum and not on scientific reasoning, inquiry and thinking skills as needed for success in science education and everyday life (Nemeth & Korom, 2012). Through basic education, learners are to develop the competencies, attitudes and values needed for full participation in society by learning to use, acquire, construct, evaluate, and transform knowledge. Learning to learn is at the core of this process, and in a knowledge based society, this continues as a lifelong learning (NIED, 2008).

2.3 Motivation to Learn Science

When measuring the motivation to learn science, we attempted to determine why students strive to learn science, what emotions they feel as they strive, how intensively they strive, and how long they strive (Glyn, Brickman, Armstrong, & Taasoobshirazi, 2011). Motivation to learn science benefits young students who aspire to be future scientists (Bryan, Glyn, & Kittleson, 2011). Furthermore, (Bryan et al., 2011) assert that "motivation to learn science does benefit all students by inculcating and improving their scientific reasoning skills and literacy, which is the ability to understand scientific knowledge, identify important scientific questions, draw evidence-based conclusions, and make decisions about how human activity affects the natural world" (p. 1050). Other studies' findings on the influence of students' attitudes towards science, learning strategies and science learning achievement (Wan, & Lee, 2017) further strengthen the significance of developing positive attitudes amongst students when they learn science. The importance of all students becoming scientifically literate cannot be overemphasized, as it is advocated for internationally. The items on the Science Motivation Questionnaire were designed to serve as empirical indicators of components of students' motivation to learn science. From the perspective of social cognitive theory, the motivation of students to learn science was examined. The students responded to the Science Motivation Questionnaire II (Glyn, et al., 2011), which assessed five motivation components: intrinsic motivation, self-determination, self-efficacy, career motivation, and grade motivation.

3. The Present Study

The current study aimed to assess students' reasoning skills through the online assessment instruments, and to explore the relationship between reasoning skills in a science context and motivation to learn science in Oshana region, Namibia, guided by four research questions:

1. How well do Grade 5 and 7 students perform on a scientific reasoning test?

2. What are the differences between scientific reasoning skills for grade 5 and grade 7 students?

3. Are there gender differences in scientific reasoning and motivation to learn science?

4. What is the relationship between scientific reasoning and motivation to learn science among the students?

4. Research Methodology

4.1 Participants

The sample for the study was drawn from the fifth- and seventh-graders (N=616; 268 boys; 348 girls; age M=12.40, SD=1.19) from five different schools in the towns of Oshakati and Ongwediva. The fifth and seventh grade learners were chosen as they are at their prime age to develop the competencies, attitudes and values needed for full participation in society by learning to use, acquire, construct, evaluate, and transform knowledge (Wan, & Lee, 2017). The sample was N=270 for Grade 5 (120 boys; 150 girls, age M=11.19, SD=.68). In Grade 7, 346 students took part in the study (150 boys; 196 girls, age M=13.23, SD=.61). The schools were selected based on the availability of an ICT infrastructure at the school. Therefore, the sample is not representative; typically, these schools are attended by students from above average social backgrounds. All the Grade 5 and 7 students that were present during the day of the test took part in the assessment. Due to the selection of the schools, the sample consisted of a number of students whose parents have attained a high level of education (Table 1). The level of parental education is considered as the best indicator for students' socio-economic background (Keller, Neumann, & Fischer, 2017). It can be attributed to this relatively homogeneous, above average social background composition of the sample that ANOVA showed no significant difference between students' achievement according to parents' level of education.

Table 1. The distribution of the level of education among the students' parents

Educational level	Mothers (%)	Fathers (%)
Didn't finish elementary school	10.2	7.0
Elementary school	5.2	8.0
Vocational school	3.0	3.7
Mature exam	5.8	3.6
Undergraduate degree	7.0	11.2
Advanced degree	37.5	35.1
I don't know	30.0	27.4
Missing data	2.0	4.0

4.2 Instruments

4.2.1 Scientific Reasoning Test

This study was based on a scientific reasoning skills test, where students were required to use their cognitive skills (scientific reasoning) to answer questions based on different sub-constructs of scientific processes. It assesses different reasoning skills essential for learning science and learning in general.

The test was originally developed by the MTA-SZTE Science Education Research Group to assess Hungarian students' reasoning skills. Items were developed based on Lawson's Classroom Test of Scientific Reasoning (Lawson, 2000) framework model. The Lawson model was simplified to accommodate the basic skills required in 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 scientific reasoning skills among primary school students up to the secondary school students.

Since the tasks were developed in a Hungarian context, they were translated by a professional translator into English in consultation with those involved in task development. Tasks that were deemed unsuited to 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 review and suggestions. Two experienced teachers from Namibia also reviewed the translated texts. The final online assessment tool consisted of 16 tasks (36 items) assessing various reasoning skills in the science context (Korom et al., 2017), such as conservation (Figure 1), proportional (Figure 2), correlational and probabilistic reasoning as well as working with logical operations (Figure 3; Korom et al., 2017). The reliability index for the scientific reasoning skills test was acceptable (Cronbach's alpha=.74) for the whole sample, a surprising result, given the fact that the instrument was newly developed (Korom et al., 2017). In Grade 5, reliability was just above the threshold (Cronbach's alpha=.69), and reliability for Grade 7 was within the acceptable range as well (Cronbach's alpha=.72). However, the test yielded lower reliability indices at the level of the subscales (ranging between (α=.40 and α=.70).

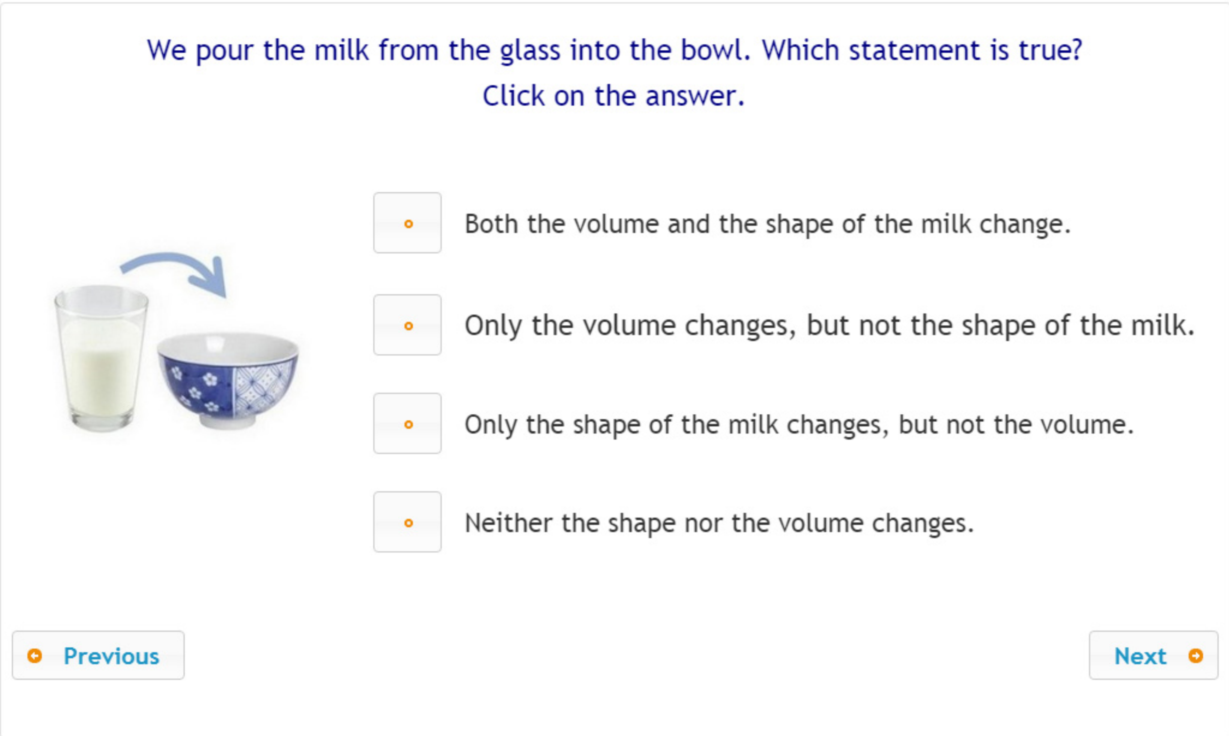

Figure 1. Sample task for conservation

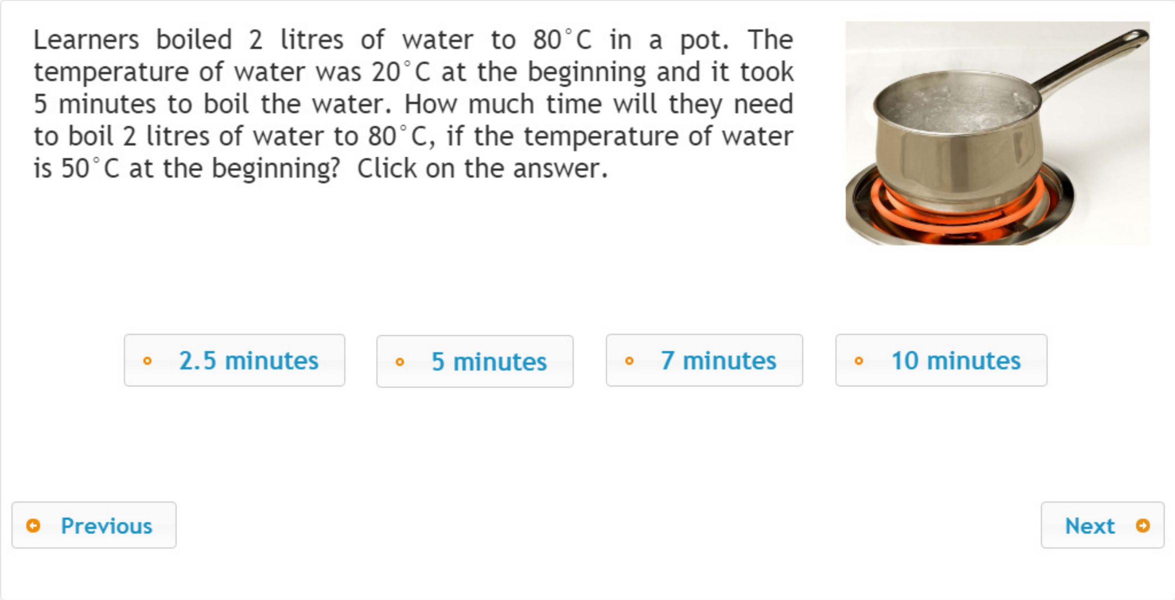

Figure 2. Sample task for proportional reasoning

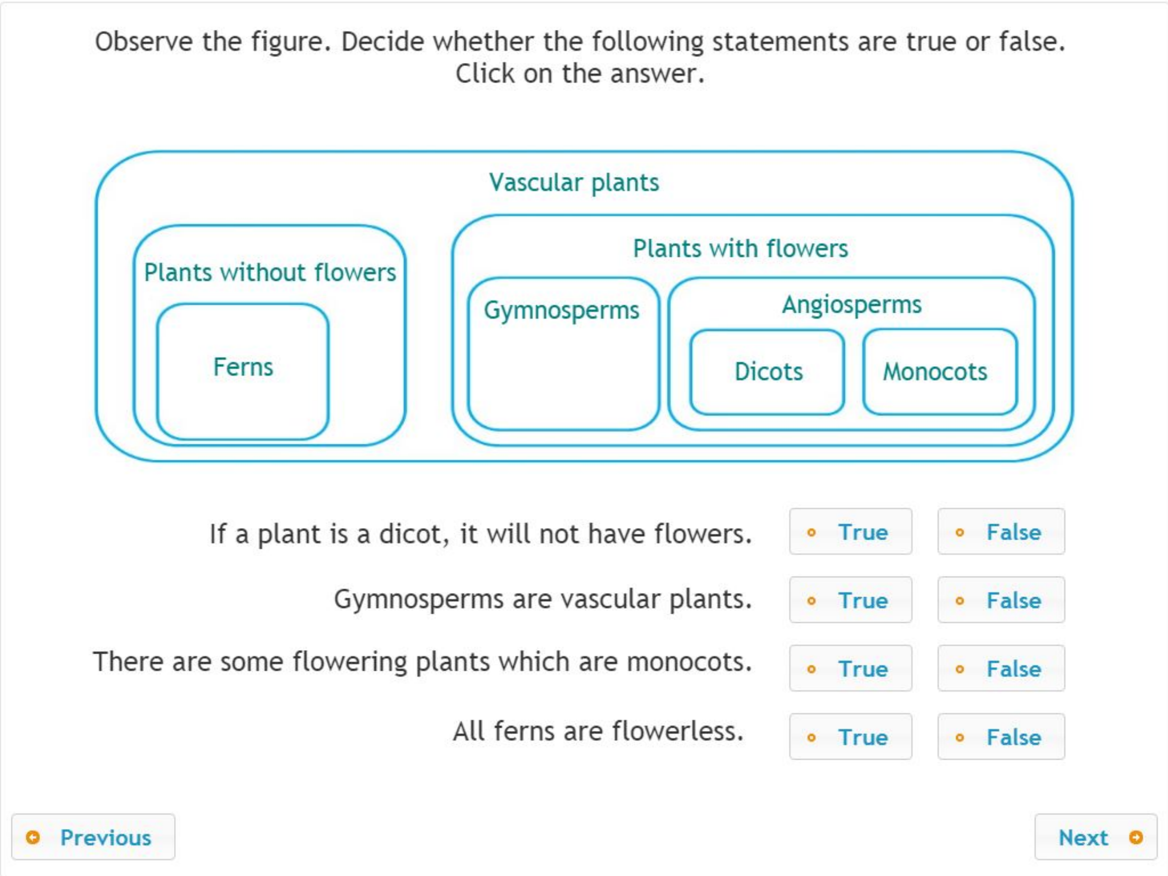

Figure 3. Sample task for logical classifications

4.2.2 Science Motivation Questionnaire II (SMQ)

The revised five point Likert scale Science Motivation Questionnaire (SMQ, Glynn et al.,

2011) was used without adaptation in the present study based on social cognitive theory. The SMQ is a paper-and-pencil instrument, which was changed to an online version tool for the purposes of this study. Glynn et al., (2011) assert that regardless of age, if students are motivated both intrinsically and extrinsically, they are likely to do better in any given tasks, hence the choice of this questionnaire.

The questionnaire was also sent to the same expert mentioned in the scientific reasoning section for quality assurance. For the science motivation questionnaire, reliability (Cronbach's alpha=.91) was quite high for the whole sample, which is not surprising, given the fact that this is a previously validated instrument available in the literature (for the subtests, it ranged between .65 and .82). In Grades 5 and 7, the reliability indices were α=.90 and α=.91, respectively. These indices match Glynn et al.'s findings (2011), when the instrument was being validated (α= .95) with the freshmen in the US.

In respect of reliability, both the scientific reasoning test and the motivation questionnaire behaved the same in both grades, as there is no significant difference between the Cronbach's alphas.

One-parameter RASCH analyses were also carried out to develop a more detailed idea of 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 (Figure 4). The analysis revealed that there were no items at a low-ability level, especially in Grade 5, and that there were several items at the top which no students were capable of doing correctly. This means that the test was somewhat difficult for the students, as the same trend continued in Grade 7 as well. However, the distribution in Grade 7 was somewhat better than that of Grade 5. Indeed, as reported in the first section (grade differences), Grade 7 performed much better than Grade 5. This can be attributed to the fact that 7^{th} graders may have improved their English language capabilities. This finding also conform with what is found in the literature that SR progress with age. Few students were at the lower end of the distribution, with few high-ability items at the top.

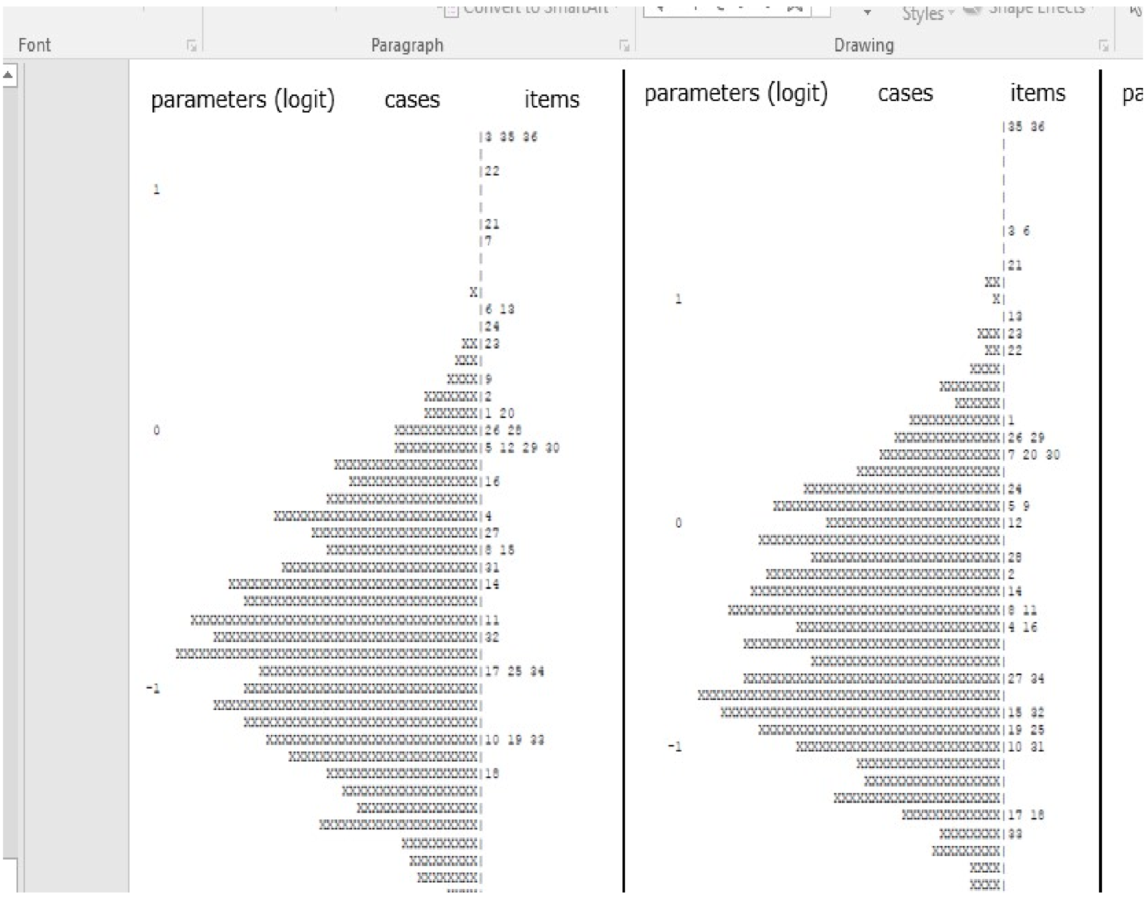

Figure 4. Person-item maps for the reasoning skills test

5. Data analysis

As indicated earlier, the purpose of this study was to assess students' abilities in scientific reasoning skills and to explore the relationship between reasoning skills with motivation to learn science in Namibia. For research questions 1, 2 and 3, independent sample t-tests were used to find the differences in performance between the grades and between genders. Log file output was analyzed qualitatively in relation to questions 1 and 2. In addition, for question 4, correlations were used to explore the relationship between the reasoning test and motivation to learn science. Analysis of variance (ANOVA) was applied to examine the differences between students' achievement and their parents' level of education (socio-economic status). Furthermore, item response models (IRT) were used as they are in line with our research goal of determining students' ability on the reasoning test as well as the extent to which item difficulty levels correspond to the abilities of the assessed students (Adams & Wu, 2002).

6. Results

6.1 Grade Differences - Developmental Tendencies

Overall, the results indicate that the scientific reasoning skills test was moderately difficult for the students, as the older age group only scored an average of: M=40.56%; SD=13.46% (see Table 2). However, they performed fairly well in the proportional and correlational subtests compared to the rest of the subtests. The distribution of the achievement revealed that none of the students achieved above 80%, and only three of the students scored below 10%. Furthermore, the results indicate that the test might be too difficult for the sampled primary students.

The difference between the two grades (Table 2) in performance on the scientific reasoning test was statistically significant ($t(616) = 7.87$, $p<.01$). This means that scientific reasoning skills developed with age, as the Grade 7 students performed much better than the Grade 5 students. The difference in standard deviation units was $d=0.52$, indicating a medium impact of the two years of education. These results could suggest the need to improve reasoning and thinking skills among students at the primary level. Apart from that, the differences in the two grades may also be attributed to the language capabilities of the students since English is the medium of instructions in Namibia. However, most students speak might not be fluent enough to understand the demand of the questions as they speak their mother tongue at home and English only in the classroom or within the school premises. As stated on the literature review section, the curriculum demands and how it is delivered, the level of exposure to computer testing by students may also have contributed to the grades performances and differences found.

Table 2. Descriptive statistics for the scientific reasoning subtests

Scales	Both grades (N=621)		Grade 5 (N=275)		Grade 7 (N=346)	
	Mean (%)	SD (%)	Mean (%)	SD (%)	Mean (%)	SD (%)
Scientific reasoning	**37.83**	**13.34**	**34.39**	**12.34**	**40.56**	**13.46**
Conservation of mass and volume	35.07	20.22	32.23	18.51	37.32	21.23
Proportional reasoning	41.01	19.81	37.13	19.64	44.09	19.43
Correlational reasoning	43.00	35.52	38.00	35.76	46.97	34.86
Probabilistic reasoning	36.67	28.05	33.36	27.49	39.31	28.26
Classification	38.44	18.61	34.75	17.55	41.37	18.94

6.2 Tasks results

The tasks analysis on conservation of mass and volume found that 70.2% of the students could not answer the question correctly, 22.3% of students thought that both the volume and the shape of the milk would change, 28.6% believed that only the volume would change, and 12.1% concluded that neither the shape nor volume would change. This result shows that these students have a serious problem with understanding the concept of volume. The same trend continues with the task on proportional reasoning; about 78% of the participants could not supply the correct answer. According to the log analysis, many who failed thought that the task entails multiplying by two or dividing by 0.5, but students mixed up the methods (40.5%). Only 19.9% of students were able to complete all four logical operation tasks correctly. It must be noted that no explicit science knowledge is necessary to do the task of logical operations and classifications. One only needs to use his/her thinking and reading comprehension skills. Thus, the low achievement also reveals fundamental obstacles in science learning as students seem to have major problems understanding basic science.

The Science Motivation Questionnaire scores were relatively high (M=3.06, SD=.71); thus, students reported that they are motivated to learn science (Table 3). However, the lowest score was found on intrinsic motivation and self-determination. This means that students' performance on science tasks was low even though they indicated that they are motivated to learn science.

Table 3. Descriptive statistics for Science Motivation Questionnaire II (SMQ)

Scale	Both grades		Grade 5		Grade 7	
	Mean (%)	SD (%)	Mean (%)	SD (%)	Mean (%)	SD (%)
SMQ	**2.92**	**.75**	**2.84**	**.76**	**3.06**	**.71**
Intrinsic motivation	2.78	.85	2.72	.85	2.85	.86
Self-efficacy	3.04	.89	2.94	.90	3.19	.85
Self determination	2.84	.86	2.76	.89	2.98	.79
Grade motivation	3.06	.84	2.97	.85	3.19	.82
Career motivation	2.96	.88	2.87	.89	3.10	.86

6.3 Gender Differences

Tables 4 and 5 indicate how students performed in both constructs in terms of gender. The results showed that no significant differences were found between genders ($p > .05$). This conforms to current research findings on scientific reasoning and science education in general that gender does not influence one's performance in science (OECD, 2016; Korom et al., 2017). No gender differences were found for the SMQ either, and this resonates well with Glynn et al.'s (2011) results.

Table 4. Gender differences on the scientific reasoning subtests

Scales	Boys (N=268)		Girls (N= 348)			
	Mean (%)	SD (%)	Mean (%)	SD (%)	t	p
Scientific reasoning	**36.75**	**13.20**	**38.75**	**13.46**	**1.84**	>.05
Conservation of mass and volume	36.34	20.18	36.51	20.21	1.93	
Proportional reasoning	40.38	20.11	41.54	19.58	.72	
Correlational reasoning	41.79	33.19	43.97	36.50	.76	
Probabilistic reasoning	36.57	26.24	37.20	28.00	.19	
Classification	37.58	18.51	39.12	18.74	1.02	

Table 5. Gender differences on the Science Motivation Questionnaire subtests

Scales	Boys (N=187)		Girls (N= 244)			
	Mean (%)	SD (%)	Mean (%)	SD (%)	t	p
SMQ	**2.92**	**.71**	**2.93**	**.78**	**.18**	
Intrinsic motivation	2.77	.84	2.79	.87	.28	
Self-efficacy	3.04	.87	3.04	.90	.10	>.05
Self determination	2.88	.82	2.81	.89	.80	
Grade motivation	2.97	.87	3.12	.82	1.82	
Career motivation	3.00	.86	2.96	.89	.05	

6.4 Relationships between Scientific Reasoning and Motivation to Learn Science

The correlation matrix in Table 6, shows bivariate relationships between the variables of the two constructs was explored in this study. Strong positive correlations were found between scientific reasoning and its subscales (r=.50-.80, p<.01). In the same vein, the subscales for the SMQ were highly correlated with each other (r=.83-.86, p<.01). However, a weak relationship was found between scientific reasoning test achievement and motivation to learn science (r=.21, p<.01). This means that even though students scored low on the scientific reasoning tests, they are still motivated to learn science. Some of the SMQ subscales did not significantly correlate with the reasoning test (e.g. intrinsic motivation, self-determination, and career motivation).

Table 6. Correlations between the tests and subtests for both age groups

Variables	1	2	3	4	5	6	7	8	9	10	11
1 Conservation											
2 Proportional r.	.27**										
3 Correlational r.	.15**	.10*									
4 Probabilistic r.	.19**	.16**	.14**								
5 Classification	.19**	.20**	.14**	.10*							
6 Reasoning test	.80**	.63**	.34**	.45**	.50**						
7 Intrinsic motivation	.11*	.11*	n.s.	.11*	n.s.	.14**					
8 Self-efficacy	.11*	.15**	.10*	n.s.	n.s.	.16**	.65**				
9 Self-determination	.14**	.11*	n.s.	n.s.	n.s.	.16**	.65**	.63**			
10 Grade motivation	.18**	.12**	.12*	n.s.	.10*	.21**	.64**	.66**	.63**		
11 Career motivation	.20**	.15**	n.s.	.10*	n.s.	.21**	.63**	.65**	.60**	.63**	
12 Motivation test	.18**	.14**	n.s.	n.s.	n.s.	.21**	.85**	.86**	.83**	.84**	.84**

**Correlation is significant at 0.05 (2-tailed), **Correlation is significant at 0.01 (2-tailed); n.s. = non-significant correlations*

7. Discussions

The online assessment instruments for SR proved to be reliable despite the low mean achievement. This suggests that the test in its present form is a better fit for an older age group; easier items should be added to assess fifth- and seventh-graders. However, the results conform to those of Bao et al. (2009), in which a scientific reasoning test by Lawson (2000) was used to compare the reasoning abilities of US and Chinese students. A similar general developmental trend was found for test scores between these two populations (Bao et al., 2009). The explanation for the low performance could be attributed to language. It is acknowledged here that English is used in Namibia as a second language, but the majority of the students only use English in school. Therefore, the English proficiency level of the students might have contributed to the low achievement factor. Additionally, the results from the study also fall within the same range as in Lawson's (1978, 2000) studies, in which he validated his instrument. Lawson's test reliability was α= .65-.75 based on the performance of eighth- to eleventh-graders in the US. This study thus confirms that a scientific reasoning test could be a good measure to assess students in Namibia from primary school to the first year of college.

The psychometric analyses of the scientific reasoning skills test and the scientific reasoning test demonstrated that the tests must be revised to carry out more reliable assessments in this age cohort. The reliability indices and the IRT analyses indicated that the test was difficult for these primary students. New items must be added with some items modified to fit the tests' difficulty to students' skills and to increase the discrimination power of the test at low skill levels. This suggests that the easiest items on the present tests may be retained for benchmarking and anchoring for international comparative studies, and even easier items should be added to fit the test to students' current developmental level. Maintaining anchor items and carrying out comparative studies may be extremely important, as the curriculum and the teaching methods should be improved to make schools more effective in developing reasoning skills. In its present form, the instrument can be used to assess reasoning and thinking skills at the secondary level.

The task analysis showed that students were unable to complete simple tasks on conservation or proportional and logical operations. This may not be attributed exclusively to the difficulty level of the items; the effort to include the twenty-first-century skills in the curriculum seems minimal. Namibian learners need to acquire twenty-first-century skills as well as students in other countries, but the education system in Namibia is not ready or equipped to impart such skills. The performance on these tasks highlights the need to train students and expose them explicitly to a different type of reasoning and thinking skills. This finding may also suggest that teacher training would be necessary on how to incorporate these skills into everyday teaching. Monitoring progress with revised instruments may provide vital feedback for developmental efforts.

Examining the relationship between reasoning skills achievement and some aspects of motivation to learn science, weak correlations were found between scientific reasoning and motivation with relatively high average scores on motivation, perhaps suggesting that regardless of the level of students' scientific reasoning skills, they still intend to learn science in general. This may provide a promising basis for science instruction in the future. Glynn et al. (2011) also reported that some components of motivation, especially self-efficacy, grade motivation, and career motivation were related to the students' college science grade point averages. However, it should be noted that the participants in these previous studies were college freshmen, who might have had clear ambitions about what they wanted to do in the future. Therefore, the results suggest that the study should be repeated with much older samples, i.e. from eighth grade to college freshmen too.

According to Figure 4, the SR showed that more items were 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 such, one may conclude that the curriculum focuses more on content and is exam-driven.

8. Conclusions

This study was one of the first attempts to carry out an online assessment in Namibia at the primary level. The results indicate that technology-based assessment may provide schools and teachers a user-friendly instrument for monitoring the development of students' thinking skills. The study confirms the technical feasibility of technology-based assessment in a country where implementation of the relevant equipment is still in an initial phase. The translation of the items developed in another educational system from another language resulted in instruments with acceptable parameters, and the results provided a firm basis for further improvement of the instruments.

It should also be noted that our sample was biased towards socially advantaged students, and one wonders what would happen if the same tasks were administered in primary schools in and around the countryside, where most students are from more modest socio-economic backgrounds. With regard to the non-significant correlations found between self-efficacy and test performance, this might be attributed to language. Self-efficacy might be a new concept for most primary students. The low correlations between reasoning skills in science and motivation and the relatively high average scores on motivation suggest that regardless of the level of their reasoning skills, students are eager to learn science in general. This may provide a promising basis for science instruction in the future.

9. Limitations and recommendations

With regard to online assessments, the technology-based assessment of reasoning skills proved to be a hurdle to overcome in the Oshana region at the moment. From the general observation, the situation in most public schools is that technology infrastructures are lacking. Moreover, this has currently become one of the major limitations of this research. A longitudinal study would demonstrate the situation on the ground concerning ICT infrastructure in schools. It is also envisaged that a paper-and-pencil test would be sufficient until politicians and the private sector come on board to improve the ICT situations in public schools. Based on the results of this pioneering study, the possibilities for developing an online evaluation system can be further explored in the near foreseeable future. Also, perhaps a paper-and-pencil test needs to be used, and the online test should be piloted with older students, those in grades higher than 5 and 7.

References

Adams, R. J., & Wu, M. L. (Eds.). (2002). *PISA 2000 technical report*. Paris: OECD Publications.

Adey, P., & Csapó, B. (2012). Developing and assessing scientific reasoning. In B. Csapó, & G. Szabó (Eds.), *Framework for diagnostic assessment of science* (pp. 17−53). Budapest: Nemzeti Tankönyvkiadó.

Bao, L., Cai, T., Koenig, K., & Fang, K. (2009). *Learning and scientific reasoning.* Washington, DC: AAAS. https://doi.org/10.1126/science.1167740

Brom, C., Šisler, V., & Slavík, R. (2009). Implementing digital game-based learning in schools: augmented learning environment of 'Europe 2045'. *Multimedia Systems*, *16*(1), 23-41. https://doi.org/10.1007/s00530-009-0174-0

Bryan, R. R., Glyn, S. M., & Kittleson, J. M. (2011). Motivation, achievement and advanced placement intent of high school students learning science. *Science Education*, *95*(6), 1049-1065. https://doi.org/10.1002/sce.20462

Crawford, B., & Cullin, M. (2005). Dynamic Assessments of Preservice Teachers' Knowledge of Models and Modelling. In K. Boersma, M. Goedhart, O. de Jong & H. Eijkelhof (Eds.), Research and the Quality of Science Education, (pp. 309-323). Dordrecht: Springer. https://doi.org/10.1007/1-4020-3673-6_25

Csapó, B., & Molnár, G. (2017). Potential for assessing dynamic problem-solving at the beginning of higher education studies. *Frontiers in Psychology*, *8*(2022). 1-12. https://doi.org/10.3389/fpsyg.2017.02022

Csapó, B., Lörincz, A. & Molnár, G. (2012). Innovative assessment technologies in educational games designed for young students. *Assessment in Game-Based Learning*, 235-254. https://doi.org/10.1007/978-1-4614-3546-4_13

Fischer, F., I. Kollar, S. Ufer, B. Sodian, H. Hussmann, R. Pekrun, B. Neuhaus, et al. (2014). Scientific Reasoning and Argumentation: Advancing an Interdisciplinary Research Agenda in Education. *Frontline Learning Research*, *2*(3), 28-45.

Glynn, S. M., Brickman, P., Armstrong, N., & Taasoobshirazi, G. (2011). Science Motivation Questionnaire II: Validation with science majors and nonscience majors. *Journal of Research in Science Teaching, 48*, 1159-1176. https://doi.org/10.1002/tea.20442

Hamers, J. H. M., de Koning, E. & Sijtsma, K. (1998). Inductive reasoning in third grade: Intervention promises and constraints. *Contemporary Educational Psychology*, *23*(2), 132-148. https://doi.org/10.1006/ceps.1998.0966

Ibrahim, B., Ding, L., Mollohan, K. N., & Stammen, A. (2016): Scientific Reasoning: Theory Evidence Coordination in Physics-based and Non-physics based Tasks. *African Journal of Research in Mathematics, Science and Technology Education, 20*(2), 93-105. https://doi.org/10.1080/10288457.2015.1108570

Iipinge, S. M., & Likando, G. N. (2012). The Educational assessment reforms in post-independence Namibia: A critical analysis. *SA-eDUC JOURNAL*, 1-10.

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. https://doi.org/10.1002/tea.21378

Korom, E., B. Németh, M., Pásztor, A., & Csapó, B. (2017). The relationship between scientific and inductive reasoning in grades 5 and 7. Paper accepted at the 17th biennial conference of the European Association for Research on Learning and Instruction (EARLI). Tampere, Finland, August 29-September 2, 2017.

Lawson, A. E. (1978). Development and validation of the classroom test of formal reasoning. *Journal of Research in Science Teaching*, 15(1), 11-24. https://doi.org/10.1002/tea.3660150103

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. https://doi.org/10.1002/tea.20172

Ministry of Education. (2007). *Education and training sector Improvement program.* Windhoek: Government Press.

Molnár, G. (2011). Playful fostering of 6- to 8-year-old students' inductive reasoning. *Thinking Skills and Creativity, 6*(2), 91-99. https://doi.org/10.1016/j.tsc.2011.05.002

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

National Institute for Educational Development (2008). *The national curriculum for basic education*. Okahandja: Ministry of Education Namibia.

NIED. (2009). *The National Curriculum for Basic Education*. Okahandja: Ministry of Education, Namibia.

Nemeth, M. B., & Korom, E. (2012). Science literacy and the Application of Scientific knowledge. In B. Csapo, & G. Szabo, *Framework for Diagnostic Assessment of Science* (pp. 55-87). Budapest: Nemzeti Tankonyvkiado Zrt.

OECD. (2007). *Assessing scientific, reading and mathematical literacy: A framework for PISA 2006*. Paris: OECD Publishing. https://doi.org/10.1787/9789264066182-ja

OECD. (2013). *PISA 2012 Results in Focus: What 15-year-olds know and what they can do with what they know*. Paris: OECD Publishing.

OECD. (2016). *PISA 2015 Results (Volume I) Excellence and Equity in Education.* Paris: OECD Publishing.

Piaget, J., & Gabain, M. (1965). *The moral judgment of the child*. New York: Free Press.

Prensky, M. (2007). *Digital game-based learning*. Saint Paul, MN: Paragon House.

Shaakumeni, S. N. & Csapó, B. (2018). A Cross-cultural Validation of Adapted Questionnaire for Assessing Motivation to Learn Science, *African Journal of Research in Mathematics, Science and Technology Education*, *22*(3), 340-350. https://doi.org/10.1080/18117295.2018.1533157

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.

Stiller, J., Hartmann, S., Mathesius, S., Straube, P., Tiemann, R., Nordmeier, V., Krüger, V., & Upmeier zu Belzen, A. (2016): Assessing scientific reasoning: a comprehensive evaluation of item features that affect item difficulty, *Assessment & Evaluation in Higher Education*, 1-13, https://doi.org/10.1080/02602938.2016.1164830

Wenning, C. J. (2006). Assessing nature of science literacy as one component of scientific literacy. *Journal of Physics Teacher Education Online*, *3*(4), 3-14.

Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review*, *27*(2), 172-223. https://doi.org/10.1016/j.dr.2006.12.001